



Instituto Hidalguense de Educación
Universidad Pedagógica Nacional-Hidalgo

Licenciatura en Intervención Educativa



Evaluación Educativa

ADVERTENCIA
ESTOS MATERIALES FUERON ELABORADOS CON FINES EXCLUSIVAMENTE
DIDÁCTICOS PARA APOYAR EL DESARROLLO CURRICULAR DEL PROGRAMA
EDUCATIVO.

TABLA DE CONTENIDOS

PRESENTACIÓN

BLOQUE I. "Constitución y problemática del campo de la evaluación"	FOLIO
➤ La evolución su conceptualización Jiménez, Jiménez Bonifacio.	1-32
➤ La evaluación. Un proceso de diálogo, comprensión y mejora. Santos, Guerra Miguel A.	33-43
➤ Evaluación: concepto, tipología y objetivos Casanova, María Antonieta.	44-62
➤ Consideraciones teóricas sobre la evaluación en la educación. Rueda, Beltrán Mario.	63-70
➤ La evaluación en el proceso educacional Lafourcade, D. Pedro.	71-78
BLOQUE II. "Los objetivos de la evolución educativa"	
➤ Ámbitos de la evaluación Casanova, María Antonieta.	79-92
➤ Patología general de la evaluación Santos, Guerra Miguel A.	93-109
➤ Evaluación de los aprendizajes Rosales, Carlos	110-124
➤ La evolución de programas para la innovación. Jiménez Jiménez Bonifacio.	125-186

BLOQUE III. "Paradigmas y técnicas de evaluación"	FOLIO
➤ Tesis para una teoría de la evaluación y sus derivaciones Díaz, Barriga Ángel.	187-212
➤ Evolución cualitativa de centros escolares. Santos, Guerra Miguel A.	211Bis- 240
➤ Nuevo enfoque para la evaluación de proyectos Sapag, Chain Nassir.	
➤ Introducción De Bono, Edward.	
➤ Portafolios del docente como responsable del desarrollo curricular y la evaluación. Giselle O, Kniep Martín.	
➤ Antecedentes del Assesment center y del abordaje a las organizaciones Schiemenson, Aldo.	
BLOQUE IV "El proyecto de evaluación"	
➤ Paradigmas y modelos de evolución. H.S. Bhola.	
➤ Técnicas e instrumentos de evaluación H.S. Bhola.	
➤ Herramienta y técnicas de la evaluación naturalista. H.S. Bhola.	

UNIVERSIDAD PEDAGÓGICA NACIONAL
PROGRAMA DE REORDENAMIENTO DE LA OFERTA EDUCATIVA DE LAS
UNIDADES UPN.

LICENCIATURA EN INTERVENCIÓN EDUCATIVA 2002.

PROGRAMA INDICATIVO

EVALUCIÓN EDUCATIVA

Área de Formación Profesional:
Básica en educación

Campo de competencia:
Instrumental.

Carácter:
Obligatorio

Créditos:
Diez

Elaboraron.
Neftali Secundino Sánchez

Unidad 122, Acapulco, Guerrero
Petrona Matus López
Unidad 203 Ixtepec Oaxaca.
Mayo de 2002

Programa indicativo
Evolución Educativa

PRESENTACIÓN

El programa Evolución educativa es un curso teórico/práctico, se impartirá durante el cuarto semestre, paralelamente con diseño Curricular, Desarrollo del Adolescente y del Adulto y, dos cursos optativos. Tienen valor curricular de 10 créditos y una carga de trabajo semanal de 6 horas.

La evolución constituye un campo de conocimiento cuya importancia ha aumentado en el transcurso de los últimos años, hasta alcanzar niveles protagónicos. Su protagonismo creciente lo ha llevado a convertirse en el centro de atención, análisis, reflexiones y debates pedagógicos y psicopedagógicos.

Esta importancia creciente responde a un sinnúmero de razones, entre otras, que la evolución constituye, uno de los espacios educativos de mayor controversia por las dudas y contradicciones que genera, respecto a las actuaciones o decisiones asociadas con ella.

Asumiendo la importancia de la evolución, como proceso inherente a la actividad educativa, el curso aporta elementos que permitirán al estudiante fundamentar la toma de decisiones para mejorar los procesos educativos en los que actúe. En efecto, el dominio conceptual y práctico de los aspectos referidos a la evaluación es fundamental para el ejercicio de la intervención educativa. En tanto arroja datos que pueden repercutir en la toma de decisiones más documentada y por lo tanto de mejorar el proceso.

De ahí la incursión en el mapa curricular de la LIE abarcando tres dimensiones. Programas educativos; sujetos, contextos y procesos de aprendizaje en el curso Planeación y evaluación Institucional ubicado en el 6. Semestre. Es decir, la evolución tienen como propósito Aportar elementos que sirvan para tomar las decisiones más acertadas sobre el sujeto o el objeto que se actúa.

En este programa partimos de reconocer que sobre la evolución hay una diversidad de concepciones y prácticas y en ese sentido, implicaciones varias, de ahí que es necesario que el licenciado en intervención educativa cuente con un programa amplio y claro de las limitaciones y posibilidades que ofrece cada uno de los enfoques de la evolución.

COMPETENCIA

Aplicar teorías, métodos y procedimientos pertinentes para evaluar elementos, procesos y actores del proceso educativo,. Orientados a sustentar la toma de decisiones para mejorar los resultados y el impacto de dicho proceso.

ORGANIZACIÓN DE LOS CONTENIDOS

El programa se ha estructurado en cuatro bloques que describimos a continuación.

El **bloque I** está centrado en el análisis conceptual de la evaluación desde una perspectiva amplia. Se revisan funciones y características, así como las implicaciones asociadas a su práctica. Se analizan, asimismo, la relación entre evaluación y calidad educativa.

En el **bloque II** proponemos a los estudiantes un conjunto de temas que los conduzcan a comprender la especificidad del objeto evaluado. Particularmente se revisan concepciones y propuestas metodológicas centradas en la evolución de sujetos, procesos y contextos, y programas educativos, se discute también la evolución de los aprendizajes.

El **bloque III** conduce al estudiante a revisar los dos grandes paradigmas en los que se inscribe los modelos y prácticas de evaluación. Se incluyen también el tema del enfoque de evolución por competencias, considerando la emergencia cada vez más notoria en distintas experiencias del diseño curricular, mismo que da sustento a la LIE. E espera que mediante el análisis puedan ubicar límites y posibilidades que ofrecen para la comprensión y valoración de los procesos y productos educativos. Además se revisan y discuten las técnicas e instrumentos que se proponen en cada uno de estos paradigmas y enfoque para la evolución.

El **bloque IV** es de carácter práctico, se ofrece como un espacio de síntesis de los v bloques anteriores donde los estudiantes tendrán que elaborar y socializar una propuesta de seguimiento o evaluación de algún programa educativo o bien alguna iniciativa relacionada con procesos de aprendizaje específicos. Se propone articular esta actividad a la elaboración de algunas de las actividades realizadas en cualquiera de los cursos paralelos que tome el estudiante, orientadas a asegurar la transferencia y el dominio de la competencia esperada en los distintos planos de desarrollo de su plan curricular. Para apoyar a la realización de esta actividad se estudia el diseño de un proyecto de evolución enfatizando implicadas en un estudio evaluativo y los aspectos que deben considerarse en su sueño.

CONTENIDOS ESTRUCTURADO EN BLOQUES.

BLOQUE I. CONSTITUCIÓN Y PROBLEMÁTICA DEL CAMPO DE LA EVALUACIÓN.

TEMAS

- 1: Evolución: un rostro polisémico
Orígenes de la evolución.
Evolución conceptual de la evolución
Concepto, tipología y objetivos.
- 2: La funciones educativas de la evaluación
- 3: Evolución, poder y control
- 4: Objetividad y subjetividad de la evolución
- 5: La evolución y la calidad educativa.

Bibliografía

ALLEN. D. (200). La evaluación del aprendizaje de los estudiantes. Una herramienta para el desarrollo profesional de los docentes. Cap. III. Contextos eficaces, México, Piados, Pp. 213-263.

BERTONI, A. La evaluación. Un enfoque para el mejoramiento de la calidad educativa.
En: TORRES, R..M., BERTONI, A., Y CELMAN, S. La Evaluación. Novedades educativas, 2000.pp. 41-62.

CARDELI, J., ROLDÁN, F., Y MORDUCHOWICZ, Evaluación de la calidad y pago de incentivos salariales, Novedades Educativas.
Sumario NO. 120, Buenos Aires México, pp. 16-17.

CASANOVA Ma. Antonieta (19898). "La evaluación actual. Antecedentes". En: La Evaluación Educativa. México: SEP. 25-40.

DE ALBA, A (...) "Hacia una reconstrucción histórico-conceptual del campo de la evaluación educativa". En: Evaluación Curricular.

UNAM-CESU, 74-75.

DIAZ-BARRIGA, F. (200), "Formación docente y educación basada en competencias.
En: Formación en competencias y certificación profesional. Pensamiento universitario. No. 91. CESU-UNAM.

DÍAZ BARRIGA, F., Y HERNÁNDEZ, G. (1998). Estrategias Docentes para un Aprendizaje Significativo. Una Interpretación Constructivista. Mc Graw Hill, México.

GARCÍA, J., MONTEMAYOR, J., Y LÓPEZ, R. (R/F). La educación basada en normas de competencias. Realidad en el Instituto Politécnico Nacional.

GAENER, H. (1998). Más allá de la Evolución: Los componentes de una EDUCACIÓN DE LAS INTELIGENCIAS Múltiples. En: Inteligencias Múltiples, la teoría s en la práctica. Paidós, México, pp. 173-196.

GONCZI, A. Enfoques de la EDUCACIÓN Basada en Competencias: la Experiencia Australiana (segunda parte). La Academia Noviembre-diciembre de 1997.

GUTIERREZ, B., RODRIGUEZ, S, (s/f). Formación Docente Basada en Competencias.

Hager, P. (1993). Conceptions of competent. Philosophy of Education Society.

LAJOIE., S., LESGOLD, A. (1992). Dynamic Assessment of Proficiency for Solving Procedural Knowledge Tasks. Educational Psychologist, 27 (3), 365-384.

LEMA, Y. (2000). La evaluación de aprendizajes en un currículo por competencias. Educación primaria al final de la década: políticas curriculares en el Perú y los países andinos. Lima: Tarea, 2000.

LÓPEZ, M. (1999). A la calidad por la evacuación. Barcelona, Praxis.

LYONS, N. (2000). (comp). El uso de portafolios. Propuestas para un nuevo profesionalismo docente. Argentina, Amorrortu.

MARTÍN-KNIEP, G (2001). Portafolios del desempleo de maestros, profesores y directivos,. La sabiduría de la práctica. Paidós, México.

OLMEDO, J. (s/f). Evaluación del aprendizaje, mecanograma.

PARLETTR, M. Y D. Hamilton (1985). "La evaluación como iluminación". En: Gimeno.

Sacristán y A. Pérez Gómez (Comps) La enseñanza: su teoría y su práctica.. Madrid: akal, 450-4654.

SALUJKA, S. (S/F). La capacitación basada en competencias en el Reino Unido.

STUFFLEBEAM, D., Y SHINFIELD, A. (12000). Evaluación sistemática. Guía teórica y práctica. Barcelona, Paidós/MEC.

BLOQUE II. LOS OBJETIVOS DE LA EVALUACIÓN EDUCATIVA.

Temas

1. ¿Qué evaluar?
2. Sujetos, procesos y contextos.
3. Evaluación de programas.
4. Evaluación de los aprendizajes.

Bibliografía

ALVAREZ, M. (2001). "La retórica de la evaluación". En *Evaluar para conocer, examinar para excluir*. Madrid, Morata, pp. 40-50.

BRUNNER, J. Evaluación de la calidad en el nuevo contexto latinoamericano. . Presentación en el Seminario Internacional "Evaluando la Evaluación" de la Red Internacional para el Aseguramiento de la Calidad en la Educación Superior (INQAAHE). Santiago de Chile, 4 de mayo 1999.

CEMBRAMOS, F. D. MONTESINOS Y MARÍA BUSTELO (1989) "La evaluación. En. *La animación sociocultural., una propuesta metodológica*. Madrid, popular.

DANIELSON, CH., Y ABRUTYN, L. (1999). Una introducción al uso de portafolios en el aula. México, FCE.

GARNER, H. (1998). Más allá de la evaluación. Los componentes de una Educación de las inteligencias Múltiples,. En: *Inteligencias Múltiples. La teoría s en la práctica*. Paidós, México, pp. 173-196.

LAJOIE, S., LESGOLD, A (1992). Dynamic Assesment of Proficiency for Solving Procedural Knowledge Tasks. *Educational Psychologist*, 27 (3), 365-384).

LYONS, N (2000). (comp). *El uso de portafolios, propuestas para un nuevo profesionalismo docente*. Argentina, Amorrortu.

MARTÍN-KNIEP, G. (2001). *Portafolios del desempeño de maestros, profesores y directivos. La sabiduría de la práctica*. Paidós, México.

OECD (1996). *Evaluating and Reforming Education Systems*. OECD, París.

OLMEDO, J. (s/f). *Evaluación del aprendizaje, mecanograma*.

PARLETT, M. Y D. Hamilton (1985). "La evaluación como iluminación" En: Gimeno Sacristán y A. Pérez Gómez (COMPS) la enseñanza: su teoría y su práctica, Madrid: akal, 450-465.

RUEDA, M., DÍAZ BARRIGA, F. (2000). (Comps). Evaluación de la docencia. Perspectiva Actual. Paidós Educador, México.

SALUJA, S (s/f). la capacitación basada en competencias en el reino Unido.

UPN-SEP. Contexto y Valoración de la Práctica Docente, 1998.

BLOQUE III, PARADIGMAS Y TÉCNICAS DE EVALUACIÓN.

Temas

1. Paradigmas y modelos de evaluación.
2. Técnicas y recursos para la evaluación en el paradigma cuantitativo. }
3. Técnicas y recursos para la evaluación en el paradigma cuantitativo.
4. Enfoque de Evaluación por Competencias: implicaciones para los programas, los docentes, los estudiantes.

Bibliografía.

BARNETT, R. (2001). Los límites de la competencia. Barcelona, Gedisa.

BHOLA, H.S. (1992). "Paradigmas y modelos de evaluación" En: La evaluación de proyectos, programas y compañías de alfabetización para el desarrollo. Santiago. De Chile. UNESCO, 27-52.

DIAZ-BARRIGA, F. (2000). "Formación docente y educación basada en competencias. En: Formación en competencias y certificación profesional. Pensamiento universitario. No. 91. CESU-UNAM.

DIAZ-BARRIGA, F. (2000). "formación docente basada en competencias. En: Formación en competencias y certificación profesional. Pensamiento universitario. No. 91. CESU-UNAM.

DÍAZ BARRIGA, F., Y HERNÁNDEZ G. (1998). Estrategias Docentes para un Aprendizaje Significativo. Una Interpretación Constructivista. Mc. Graw Hill, México.

GARDNER, H. (1998). Más allá de la Evaluación: Los componentes de una Educación de las inteligencias Múltiples En: Inteligencias Múltiples, la teoría en la práctica. Paidós, México. Pp. 173-196.

GONCZI, A. Enfoques de la EDUCACIÓN Basada En Competencias: la Experiencia Australiana (segunda parte). La Academia. Noviembre-diciembre de 1997.

GUTIERRES, B., RODRÍGUEZ, S. (S/F). Formación Docente Basada en Competencias. HAGER, P. (1993). Conceptions of competente. Philosophy of Educations Society. LAJOIE, S., LESGOLD, A. (1992). Dynamic Assesment of Proficiency for Solving Procedural Knowledge Tasks, educational Psychologist, 27(3), 3655-384.

LEMA, Y. (2000). La evaluación de aprendizajes en un currículo por competencias. Educación primaria al final de la década. Políticas curriculares en el Perú y los países andinos. Lima: Tarea, 2000.

LYONS, N (2000). (COmp). El uso de portafolios. Propuestas para un nuevo profesionalismo docente. Argentina, Amoirrortu.

MARTÍN-KNIEP, G. (2001). Portafolios del desempeño de maestros, profesores y directivos. La sabiduría de la práctica. Paidós, México,

OLMEDO, J. (S/F). Evaluación del aprendizaje. Mecnograma.

PARLETT, M. Y D. Hamilton (1985) "La evaluación como iluminación". En: Gimeno Sacristán y A. Pérez Gómez (Comps.) La enseñanza: su teoría y su práctica,. Madrid: akal, 450-4654.

SCHIMELKES, S. (1996). La evaluación de los centros escolares. Potencia presentada en el Taller sobre Evaluación de Docentes y de Centros Educativos, organizado por el Programa de medios e Instrumentos para la Evaluación de la Calidad de la Educación de la Organización de ESTADOS Iberoamericanos para la Educación, la Ciencia y la Cultura y por la secretaria de Educación Pública de México, en Cancún, Q.R., México, del 11 al 13 de junio de 1996.

SALUJA, S. (S/F). La capacitación basada en competencias en el Reino Unido.

UPN. (2000) Programa de reordenamiento de la Oferta Educativa de las Unidades UPN. Licenciatura en Intervención Educativa 20002. UPN, México.

BLOQUE IV. EL PROYECTO DE EVALUACIÓN.

Temas

1. El proyecto de evaluación.
 - a) Elección y caracterización del objeto.
 - b) La construcción de los referentes de la evaluación.
 - c) Construcción de categorías, criterios, parámetros, indicadores.
 - d) Diseño de instrumentos.
 - e) Trabajo de campo.
 - f) Sistematización de la información.
 - g) Resultados.
 - h) Interpretación
 - i) Conclusiones.
 - j) El informe de evaluación.

Bibliografía.

BHOLA, H.S. (1992) "el proceso en una mirada. Herramientas y técnicas de la evaluación racionista". En: La evaluación de proyectos, programas y campañas de alfabetización para el desarrollo. Santiago de Chile, UNESCO, 232-242 Y 162-198.

BHOLAS, H.S. (1992). "El proceso en una mirada: herramientas y técnicas de la evaluación racionista". En: La evaluación de proyectos programas y campañas de alfabetización para el desarrollo, santiago de chile. UNESCO, 2342-242 Y 168-198.
LEITE, E. Evaluando el proyecto institucional. Novedades Educativas. Sumario No. 120, Buenos Aires México. P. 14.

RUEDA, M., DÍASZ BARRIGA, F. (2000). (Comps.). Evaluación de la Docencia. Perspectiva presentada en el Taller sobre Evaluación de Docentes y de Centros Educativos, organizado por el Programa de medios e Instrumentos para la Evaluación de la Calidad de la Educación de la Organización de ESRTADOS Iberoamericanos para la Educación, las Ciencia y la Cultura y por la Secretaria de Educación Pública de México, en Cancún, Q.R., México, de 11 al 13 de 1996.

UPN. (2000) Programa de Reordenamiento de la OFERTA DE LAS unidades
UPN. Licenciatura en Intervención Educativa 2002. UPN, México.

SUGERENCIAS METODOLÓGICAS

La coordinación del curso por parte del maestro es un factor fundamental. Esta debe estar orientada a crear un ambiente de aprendizaje que promueva la discusión grupal, la participación activa del estudiante, repleción, análisis, así como la expresión y oral y escrita, asimismo permitirá que la comunidad de participante establezca los mecanismos más ricos para compartir verdaderas experiencias de aprendizaje significativos. Los **saberes teóricos** es posible y viable abordarlos con actividades variadas (trabajamos escritos, revisiones bibliográficas, seminarios de discusión, presentaciones en foros estudiantiles, encuentros con especialistas, etc.) que permitan reconocer el campo de la evolución educativa. Las competencias procedimentales podrán ser trabajadas elaborando un proyecto de evolución en el que el estudiante a partir del enfoque o paradigma teórico del que parte, incluya todos los componentes. En el caso de saber ser y convivir se sugiere al docente, organice el desarrollo de las actividades, buscando que los estudiantes intercambien distintos puntos de vista., establezcan diálogos acerca de los problemas presentes en el panorama educativo actual, propongan estrategias de solución de problemas y arriben a toma de decisiones conjuntas.

CRITERIOS DE EVOLUCIÓN.

Tomando como referencia la competencia del curso, los contenidos y los alcances descritos en los y bloques, se sugiere al docente obtenga información mediante diferentes técnicas y procedimientos para identificar el desarrollo de las competencias demandada por el curso. Entre otros, puede solicitar al estudiante trabajos en mapas conceptuales, diagramas, realización y evolución de proyectos. Organización de eventos y encuentros estudiantiles, exposición y defensa de trabajos e ante colegas, etc. El portafolios de desempeño en este caso se considera una herramienta útil que permite al docente ir colectando testimonios de aprendizaje significativos, de distinto tipo y durante todo el proceso.

Reconocer también las búsquedas que el estudiante haya al avanzar por su propia cuenta en el dominio del curso. En todo caso las propuestas deben conducir a obtener evidencias a lo largo de todo el proceso, y al final del mismo, del cumplimiento de nivel de desempeño señalado. el conjunto de estrategias docentes, procedimientos aplicados., instrumentos empleados y las distintas situaciones creadas por los docentes, deben asegurar la participación activa del estudiante, y haga evidente el dominio de los contenidos y competencias diversas promovidas en el curso.

El curso culmina con la elaboración de un proyecto de evaluación en el que el estudiante sintetice las competencias generales alcanzadas, considerando los

distintos elementos según el paradigma o enfoque elegido. El proyecto elaborado deberá incluir, entre otros, los siguientes elementos:

Nombre y/o tipo de objeto a evaluar

Perspectiva elegida

Determinación del tipo de programa, población y objeto a evaluar

Metodologías y técnicas área el desarrollo del proyecto

Instrumentos diseñados o aplicados para la realización del proyecto

Resultados

Conclusiones

BIBLIOGRAFÍA GENERAL

BARNETT, R. (2001). Los límites de la competencia. Barcelona, Gedisa.

BHOLA, H.S. (1992). "El proceso d en una mirada: herramientas y técnicas de la evaluación racista". En: La evaluación de proyectos, programas y campañas de alfabetización para el desarrollo. Santiago de Chile. UNESCO, 2342-242 Y 168-198.

BHOLA, H.S. (1992) "Paradigmas y modelos de evaluación" En: La evaluación de proyectos, programas y campañas de alfabetización para el desarrollo. Santiago de Chile. UNESCO, 27-52.

BERUNNER, J. Evaluación de la calidad en el nuevo contexto latinoamericano. Presentación en el Seminario Internacional "Evaluando la Evaluación" de la red Internacional para el Aseguramiento de la Calidad en la Educación Superior (INQAAHE) SANTIAGO DE Chile, 4 de mayo 1999.

CASANOVA M. Antonia (1998). "La evaluación en el momento actual,. Antecedentes". En: La Evaluación educativa, México: SEI 25-40.

CEMBRANOS, F. D. MONTESINOS Y MARÍA BUSTELO (1989) "La evaluación sociocultural. Una propuesta metodológica. Madrid, popular.

DANIELSON, CH., Y ABRUTYN, L. (1999). Una introducción al uso de portafolios en el aula, México, FCE.

DE ALBA, A. (...) "Hacia una reconstrucción histórico-conceptual del campo de la evaluación educativa". En: Evaluación Curricular.

UNAM-CESU, 74-75.

DÍAZ-BARRIGA, F. (2000). "formación y educación basada en competencias. En: Formación en competencias y certificación profesional. Pensamiento universitario. No. 91. CESU-UNAM.

DÍAZ BARRIGA, F., Y HERNÁNDEZ, G. (1998). Estrategias Docentes para un Aprendizaje Significativo. Una Interpretación Constructivista. Mc. Graw Hill, México.
GARCÍA, J. MONTEMAYOR, J., Y LIPEZ, R. (S/F). La educación basada en normas de competencia. Realidad en el Instituto politécnico Nacional.

GARDNER, H. (1998). Más allá de la Evaluación. Los componentes de una Educación de las inteligencias Múltiples. En: Inteligencias Múltiples. La teoría en la práctica. Paidós, México. Pp. 173-196.

GONCZI, A. Enfoques de la educación Basada en competencias. LA Experiencia australiana (segunda parte). La academia. Noviembre-diciembre de 1997.

GUTIERRES, B., RODRÍGUEZ, S. (S/F). Formación Docente Basada en Competencias.

HAGER, P. (1993). Conceptions of competente, Philosophy of Educación Society.

LAJOIE, S., LESGOLD, A. (1992). Dynamic Assessment of Proficiency for Solving Procedural Knowledge tasks. Educational Psychologist, 27 (3), 365-384.

LEMA, Y. (2000). La evaluación de aprendizajes en un currículo por competencias, educación primaria al final de la década. Políticas curriculares en el Perú y los países andinos. Lima: Tarea, 2000.

LYONS, N. (2000). (Comp). El uso de portafolios. Propuestas para un nuevo profesionalismo docente,. Argentina, Amorrortu.

MARTÍN-KNIEP, G. (2001). Portafolios del desempeño de maestros, profesores y directivos. La sabiduría de la práctica. Paidós, México.

OECD (1996). Evaluating and Reforming Educations Systems. OECD, París.

OLMEDO, J. (S/F). Evaluación del aprendizaje. Mecanograma.

PACHECO, T., Y DÍAZ, A. (2000). Evaluación Académica. CESU-UNAM, FCE, México.

PARLETT, M, Y D. Hamilton (1985) "La evaluación como iluminación". En: Gimeo Sacristán y A. Pérez Gómez (Comps.) La enseñanza: su teoría y su práctica. Madrid,, akal, 450-465.

RUEDA, M., Y DÍAZ BARRIGA, F. (2000). (COMPS.) Evaluación de la docencia, Perspectivas Actuales. Paidós Educador, México.

SCHMELKES, S. (1996). La evaluación de los centros escolares. Ponencia presentada en el taller sobre Evaluación de Docentes y de centros educativos, organizado por el programa de medios e Instrumentos para la Evaluación de la Calidad de la educación de la Organización de Estados Iberoamericanos para la EDUCACIÓN, LA Ciencia y la Cultura y por la Secretaría de educación Pública de México, en Cancún, Q.R. del 11 al 13 de junio de 1996.

SALUJA, M S. (s/f). la capacitación basada en competencias en el reino Unido.

STUFFLEBEAM, D., Y SHINFIELD, A. (2000). Evaluación sistemática. Guía teórica Y Práctica. Barcelona, Paidós/MEC.

UPN. (2000) Programa de reordenamiento de la OFERTA Educativa de las Unidades

UPN. Licenciatura en Intervención Educativa 2002. UPN, México.

Bloque I
"Constitución y problemática del campo de la
evolución"

JIMÉNES, Jiménez, Bonifacio. "La evaluación su conceptualización" En: Evaluación de programas, centros y profesores. Madrid. Ed. DOE. Pp. 25-56.

1

**La evolución:
Su conceptualización**

I.I. Introducción

La evolución es una de las actividades más necesarias en la planificación y gestión de la formación en nuestras instituciones educativas, tanto en el ámbito formal como en el no formal. Se puede evidenciar un gran desarrollo teórico sobre la misma, aunque no puede decirse así de su práctica. Partimos de la consideración de la teoría y prácticas formativas, en cualquiera de sus quehaceres, indefectiblemente unidas y máxime cuando la evaluación refleja la práctica como tal, siendo la misma una dimensión implícita de la intervención educativa.

No obstante, hoy se verifica en la práctica de la formación de una reactualización del tema en aras a una mejora de la práctica evaluativo. Varias razones explicarían ese nuevo afloramiento. Podemos considerar, en este sentido, la propia evolución técnica y tecnológica, la nueva organización del trabajo, la conciencia de las instituciones de formación sobre la importancia del mantenimiento y desarrollo de las competencias asignadas, el presupuesto de formación que igualmente ha de ser justificado, además de que la formación es considerada como un proceso que va desde el análisis de las necesidades a la evolución. Sin olvidar los grandes beneficios para nuestras instituciones de formación y gestión con la introducción de conocimientos y procedimientos rigurosos y sistemáticos de evolución.

Es necesaria, pues, la reflexión teórica, nacida de la realidad, en aras a posibilitar luz y herramientas para la práctica evaluativo con sentido, racionalidad y funcionalidad desde una perspectiva holística de la misma. Del significado de la evolución en todas sus dimensiones, de su consideración como proceso y de la exigencia de la planificación de la misma arrancan suficientes elementos para mejorar nuestras prácticas evaluativos.

Afirmar que la evaluación es un proceso susceptible de planificación no es nada nuevo. Y esto cobra más importancia ante la evaluación de programas, por cuanto en el mismo., desde su definición, diseño, ejecución y resultados, hay que hilvanar los distintos tipos de evolución para poder ir dándole vida.

Pero también, nuestra afirmación "la evolución es un proceso susceptible de planificación" significa y comporta que hemos de: a) organizar los distintos elementos afectados, b) sistematizar las fases en su desarrollo, c) temporalizar las secuencias que puedan planificarse y d) proveer de los recursos necesarios para la misma pueda llevarse a cabo.

Todo ello supone, además de vencer las resistencias y obstáculos que pudieran encontrarse en su planificación y desarrollo, prever de un organigrama evaluativo (asignación de tareas y responsabilidades, competencia técnica), propiciar los instrumentos, técnicas y metodologías idóneas en cada momento, etc.

Desde este planteamiento es como abordamos el desarrollo del presente capítulo, que pretende centrar la evolución, su proceso, área concluir analizando los distintos tipos de evolución, sus implicaciones y finalidades, así como su exigencia metodológica e instrumental, sin olvidar los agentes y los distintos momentos evaluativos dentro de un modelo de evolución integrador.

I.1. Evolución histórica de la evolución.

Resulta ineludible revisar los antecedentes previos de la evolución para poder disponer de una idea clara sobre su conceptualización actual. Esta tarea, pues, necesaria, no resulta fácil a la luz de la proliferación de información disponible sobre el particular, y más aún, no siempre coincidente. No se trata de realizar, en el contexto de este trabajo, y estudio histórico propiamente dicho sobre evolución, sino sencillamente ilustrar tal evolución a partir de los hitos más significativos con incidencia directa en la concepción evaluativa sobre los procesos de enseñanza-aprendizaje, con repercusiones directas en la práctica de la evaluación.

Cabe significar, igualmente, que disponemos en este momento de síntesis bastante ilustrativas sobre la historia de la evolución (Guba y Lincoln, 1982 a; Cabrera y Espín,, 1986; Stufflebeam y Shinkfield, 1987; Garanto, 1989; Salvador, 1992; por citar algunos ejemplos). Si bien no existe univocidad a la hora de articular los periodos o épocas históricas evaluativas, sí que existe conciencia a la hora de significar que el fenómeno evaluativo en educación apenas cuenta con un siglo de tradición. De ahí, también, que podamos considerar que la evolución como tal es joven, a la par que o se puede justificar la falta de cultura y tradición evaluativa

en nuestros contextos de actuación práctica , como limitación o dificultad para diseñarla-ejecutarla con mayores garantías de calidad.

Sea como fuere, podemos considerar en consonancia con los autores referidos, al menos cuatro épocas clave como hitos históricos de la evolución de la evolución, especulármele a partir de la consideración de Tyler como padre de la evolución educativa. A saber: época pretyleriana, época tyleriana, época del realismo y época del profesionalismo.

I. 2. 1. Época pretyleriana

Si bien hemos apuntado que la evolución educativa propiamente dicha apenas cuenta con un siglo de historia, también cabe significar que existen antecedentes más lejanos. De hecho. Los primeros testimonios despuntan antes del año 2000 a. C., cuando los chinos ya disponían de los primeros exámenes para el acceso a los servicios públicos; incluso el propio Sócrates, en Grecia, contaba con cuestionarios evaluativos como parte de su metodología didáctica,. También existen evidencias en el Renacimiento.

De todas maneras, no se considera realmente el hecho evaluativo como tal hasta finales del siglo pasado, y en él hay que ubicarse para arrancar con un planteamiento sistemático y serio sobre la evolución educativa, sobre todo en conexión con la medida en ciencias sociales. Queremos indicar con ello que más allá de los intentos de las primeras mediciones sobre aspectos psicofísicos (capacidades sensoriales, tiempos de reacción, etc.) existe un salto al ámbito educativo en conexión con la evolución (medida) sobre pruebas mentales y físicas de los alumnos que ingresan en el Columbia Collage por parte de Galton.

Estos inicios dan pie a la sistematización y estandarización de pruebas, concretamente en relación con la inteligencia y con la ortografía en el contexto escolar (Bidet en 1904 y Rice en 1898, respectivamente). A partir de este momento surge un sinnúmero de instrumentos escolares de medida: caligrafía de Thorndikre en 1910; pruebas de escritura y ortografía de Ayres en 1918; escala de redacción de Yates y Freeman, etc.).

Más allá de estas ilustraciones, es de resaltar, en consonancia con Guba y Lincoln (1982^a), que esta época se caracteriza básicamente por lo siguiente:

- Evaluación y medición son conceptos virtualmente intercambiables.
- Evaluación y medición están insertos en el paradigma positivista propio de las ciencias físico-naturales.
- Evolución y medición se centran en la determinación de las diferencias individuales en la línea de los trabajos de Galton y Wundt.

- Evolución y medición se orienta hacia medidas estandarizadas fundamentadas en las condiciones de aplicación y los grupos normativos de referencia.
- Evolución y medición no tienen nada que ver con los programas escolares o el desarrollo del currículum, sino que se preocupan por proporcionar información sobre los sujetos.
- Medición y evaluación son concebidas en el mismo contexto que la prevaleciente metáfora industrial asumida por la escuela de la época.

I. 2. 2. Época tyleriana

A partir de los años veinte de este siglo surge un movimiento renovador en la escuela caracterizado por una aproximación más científica y sistemática al desarrollo del currículum. Dentro de esta dinámica es en los años treinta cuando ve luz una investigación de R. W. Tyler. *Estudio de los ocho años* editada en 1932, sobre alumnos formados o en escuelas progresistas de secundaria frente a escuelas convencionales y su funcionamiento posterior en la universidad. Este autor como responsable del proyecto pretendía examinar la efectividad de cierto currículum renovador y las estrategias didácticas en su desarrollo. Más allá del impacto de dicho estudio, lo relevante es el planteamiento sobre el currículum que necesitaba organizarse en torno a los objetivos educativos como base de la planificación y desarrollo, a la par que se constituyen en el referente evaluativo.

Así pues, lo genuino desde el punto de vista de la evolución también es la consideración de los objetivos, de manera que ésta no es más que el "proceso de determinar hasta qué punto han sido alcanzados realmente los objetivos educativos" (Tyler., 1950).

Conceptualmente se evidencia un cambio cualitativo importante por cuanto del planteamiento de Tyler se deriva toda una serie de implicaciones que pueden caracterizarse sintéticamente por lo que sigue:

- Es la primera vez que se hace referencia a que la evaluación es un proceso.
- La evolución se diferencia de medición. Evaluar no es simplemente aplicar un instrumento de medida o recoger información, sino también supone valor sobre la información recogida.
- Los objetivos, previamente establecidos, serán el criterio referencial para poder emitir el juicio de valor.
- Los organizadores del proceso instructivo y evaluativo son los objetivos definidos en términos de conducta, en el sentido de que son guía para la selección de contenidos, elaboración de materiales, estrategias didácticas, así como para la elaboración de instrumentos de evolución.
- Se apunta, pues, todo un modelo de evolución (que analizaremos posteriormente).

Con ello, en la línea de Tyler, la evolución proporciona un medio para el continuo perfeccionamiento de un programa, incluso para entender los logros de los alumnos con el consiguiente incremento de efectividad en las instituciones educativas. No cabe duda del nuevo significado y relevancia de la evaluación en la nueva dinámica de orientación para la mejora del currículum.

Cabe significar que los trabajos de Tyler no tienen una continuidad inmediata y de hecho no es hasta los años sesenta cuando se retoman. Con ellos queremos significar que durante los años cuarenta y cincuenta el desarrollo evaluativo no es muy importante por cuanto, en la línea de Stufflebeam y Shinkfield (1987), estos momentos se caracterizan por lo siguiente:

- La extensión de la educación mientras que decae el interés por la evolución del sistema educativo.
- Retraso en el desarrollo de los aspectos técnicos de la evolución.
- Considerable desarrollo de los instrumentos y estrategias aplicables a los diferentes métodos evaluativos como test a nivel nacional, algoritmos para designar objetivos de comportamiento, taxonomía de objetivos, etc.
- Las evoluciones dependen de las directrices educacionales dependientes de las directrices escolares locales, organizadas y financiadas por fundaciones y organizaciones profesionales.

I. 2. 3. Época del realismo

Todo lo anterior da lugar a un cierto descontento con la situación educativa americana del momento al tener que justificar la bondad de los programas educativos y su impacto en la realidad social,. Propiciándose todo un movimiento (*accountability*) de rendición de cuentas sobre el mérito y valor de dichos programas. En cualquier caso, la evolución de programas cobra una importancia significativa, sobre todo desde la óptica de la investigación evaluativa.

Superemos, pues, la evolución de los alumnos para incidir directamente más en el resto de factores que inciden en el programa educativo: profesores, centros, contenidos, experiencias de aprendizaje. Por tanto, se muestran insuficientes las pruebas de rendimiento para evaluar el programa y los objetivos como la clave articuladora del proceso de evolución.

Esto obliga a nuevos planteamientos y conceptualizaciones de la propia evolución. De hecho surge una serie de pensadores que centran su atención en el proceso, poniendo las bases del moderno concepto de evolución. Destacamos fundamentalmente dos de ellos. De uno lado, Cronbach (1963), que a partir de una revisión realizada al planteamiento tyleriano nos presenta una propuesta que pasa por la consideración de que no son los objetivos previamente establecidos el

punto de partida, sino el tipo de decisión a la que la evolución pretende servir. De manera específica se concretan en torno a los siguientes puntos:

- Si la evaluación aspira a ser un instrumento de gran utilidad para los creadores de los nuevos cursos, deberá focalizarse alrededor de las decisiones que estos creadores tendrán que tomar al desarrollarlos.
- La evolución intentará actuar durante el propio proceso de desarrollo del curso, sin esperar a que éste haya finalizado.
- La evolución se centrará en el estudio de las características estructuradas del propio programa que en estudios de corte comparativo.

Por su parte, Seriven (1967), ubicado también en la lógica de las decisiones, nos las clasifica en tres ámbitos: a) la mejora del curso (decidir qué materiales de instrucción y qué métodos son satisfactorios y en cuáles es necesario el cambio); b) la institución, selección y agrupación, etc.) y c) regulación administrativa (juzgar la calidad del sistema escolar, la de los profesores etc.). Pero, además nos realiza otras aportaciones conceptuales interesantes al clarificar entre funciones y objetivos de la evolución. En este sentido, nos indica que los objetivos siempre permanecen invariables y suponen la estimación del mérito o valor de un producto. Las funciones hacen referencia a la utilización de la información y la consecuente toma de decisiones, distinguiendo entre evolución formativa y evolución sumativa. La primera se lleva a cabo durante el proceso con el propósito de introducir las modificaciones necesarias para mejorar el programa. La segunda se realiza al final del proceso para verificar los resultados obtenidos.

1. 2. 4. Época del profesionalismo

Los trabajos de Cronbach y Seriven dan pie a una nueva época (a partir de los años sesenta en los Estados Unidos) que se caracteriza por la proliferación de modelos evaluativos y el profesionalismo en la actividad evaluadora. Todo ello en la línea de abundar en una evolución más sistemática a la par que brinda un amplio repertorio de planes evaluativos donde elegir, de gran valor heurístico y orientador. Aparecen nuevos posicionamientos paradigmáticos (evolución cuantitativa, evolución cualitativa) con nuevas herramientas, técnicas e instrumentos de evaluación.

Dado que posteriormente vamos a reparar en algunos de los modelos más relevantes, eludimos en este apartado cualquiera precisión sobre los mismos y su precursores. Sencillamente significados en esta época.

- Desplazamiento de la evolución centrada en objetivos hacia otra orientación a la toma de decisiones.
- Recrudescimiento de las críticas al paradigma científico positivista.

- Apertura a nuevos paradigmas (de corte más cualitativo con todas sus implicaciones).

I.3. En torno al concepto de evaluación

El concepto de evolución ha estado sujeto a los variantes históricos propios del ámbito educativo, tanto en su teoría como en su práctica, al y como hemos comprobado en el apartado anterior. Es decir, si revisamos la literatura en torno a evolución, podemos apreciar que se han formulado tantas definiciones como teorías y estudios se han preocupado de ello. Algunas son coincidentes, otras presentan matrices diferenciales importantes y otras tantas son eléctricas o aglutinadoras. Dichas diferencias, insistimos, obedecen a las distintas posiciones filosóficas. Epistemológicas y metodológicas que en el devenir de la reflexión-intervención educativa han predominado en un momento u otro.

No pretendemos realizar una revisión exhaustiva de la literatura en este ámbito, pero sí significar algunos hitos evolutivos, más allá de la propia historia (fases) de la evolución educativa, de cara a poder entender mejor las posturas actuales de la conceptualización de este fenómeno, de carácter más pluralista en contra del monolitismo originario. Para ello utilizaremos algunas definiciones significativas:

- a) Definiciones que contienen especificaciones o normas sobre los atributos que han de haber conseguido los alumnos después de un proceso instructivo –objetivos-.
 - Tyler (1980:69): "Proceso que determina hasta qué punto se han conseguido los objetivos educativos".
 - Lafourcade (1972: 21): "Etapa del proceso educacional que tienen por fin controlar de modo sistemático en qué medida se han logrado los resultados previstos en los objetivos que se hubieran especificado con antelación".
 - Páez (1975: 20): "Acto de comparar una medida con un estándar y emitir un juicio basado en la comparación".
 - Blom y otros (1971): "Reunión sistemática de evidencias a fin de determinar si en realidad se producen ciertos cambios en los alumnos y establecer el grado de cambio en cada estudiante".
 - Dichas definiciones, muy entroncadas con la evolución de alumnos, unas de manera explícita y otras implícita, destaca sobre todo la preocupación por el logro de los objetivos –resultados- como el objeto fundamental de la evaluación en un acercamiento puntual al final de los procesos formativos (Cierta saturación de evaluación sumativa).
- b) Definiciones que centran la evolución en la determinación del mérito o valor.
 - Suchman (1967): "Proceso de emitir juicios de valor".

- Scriven (i)&/) "Procesos por el que se determina el mérito o valor de alguna cosa".
- Nuevo (i)(#: 124): "Apreciación o juicio en cuanto a la calidad o valor de un objeto".
- Joint Comitte (i988): "La evolución es el enjuiciamiento sistemático de la valía o mérito de un objeto".

Este nuevo acercamiento comporta algo más que la comparación con estándares, que conllevan la importancia de la medida y todo el proceso que se genera en torno a la misma. Evaluar, además de medir, implica emisión de un juicio sobre la medida -valoración. Otro problema distinto será la fuente de criterios para calibrar el mérito o el valor de algo. En cualquier caso, sea como fuerte, lo cierto es que este planteamiento permite hacer susceptibles de evolución más variables del proceso formativo y abrir el mismo más allá del logro de objetivos.

- c) Definiciones que afrontan la evolución como el proceso que proporciona información para la toma de decisiones.
- Cronbach (1963: 244): "Proceso de recopilación y utilización de la información para tomar decisiones".
 - Stufflebeam y Shinkfield (1987: 183): "Proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y el mérito de la metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos impactados".
 - Fernández (1993: 11): "Emitir un juicio valorativo sobre una realidad educativa en función de unos datos y con el propósito de tomar decisiones al respecto".

Además de incluir múltiples objetos de evolución, este planteamiento que vienen a completar los anteriores, posibilita incorporar nuevas funciones y momentos evaluativos y darles pleno sentido (evolución diagnóstica, evolución formativa y evolución sumativa). Se preocupa igualmente por la utilidad de la evolución en la línea de perfeccionamiento o mejora de la práctica, con independencia de quién toma la decisión (docente, gestores educativos, políticos, etc.). También es destacable el papel de los agentes evaluadores, la relación entre de los distintos profesionales (evaluación interna. evaluación externa, auto-herteroevaluación).

- d) Definiciones que intentan sintetizar y aglutinar los aspectos anteriores.

- Provus (1971): "Comparación de las ejecuciones o comportamientos de los alumnos con ciertas normas para determinar si se debe llevar a cabo, se continúan o se concluyen en el proceso de enseñanza".

- Tenbrick (1984: 19): "Proceso de obtener información y usarla para formar juicios que a su vez se utilizarán en la toma de decisiones".
- Pérez Gómez (1983b: 431): "Proceso de recogida y provisión de evidencias, sobre el funcionamiento y evolución de la vida en el aula, en base a las cuales se forman decisiones sobre la posibilidad, efectividad y valor educativo del currículum".
- Pérez Juste (1986: 31): "Acto de valorar una realidad, formando parte de un proceso cuyos momentos previos son los de fijación de las características de la realidad a valorar, u de recogida de información sobre las mismas, y cuyas etapas posteriores son la información y la toma de decisiones en función del juicio de valor emitido".

En la medida en que las definiciones son aglutinadoras de los planteamientos previamente expuestos, hay que asumir todo lo ya dicho. Ello conlleva que conceptualicemos la evolución como.

- *Un procedimiento de recogida de información*, no improvisado, necesitado de organizar sus elementos, sistematizar sus fases, temporalizar sus secuencias, proveer los recursos, construir o seleccionar los instrumentos, etc. En cualquier caso, desde planteamientos multivariados en cuanto a los instrumentos, técnicas y métodos, así como agentes.
- *Implican un juicio de valor*, que significa que no basta recoger sistemáticamente la información, sino que ésta ha de valorarse, explicando su bondad. Téngase en cuenta que la adjudicación de un valor no significa tomar decisiones, por cuanto los evaluadores pueden realizar esta tarea y otras personas, ajeos a ellos, pueden tomar las decisiones. Queremos decir que la función principal del evaluador está en la valoración, pero no necesariamente en la toma de decisiones, que pueden corresponder a los responsables del programa i de los objetivos a evaluar. Políticos, u otros.
- *Orientada hacia la toma de decisiones*. El proceso evaluativo ha de tener una utilidad, en este sentido apuntamos entre otros a la toma de decisiones orientada ala mejora de la práctica. Esto significa, además, que la evolución ha de ser un medio pero no un fin en sí misma.

I. 4. Dimensiones de la evaluación

Aplicar la evolución exige reparar en múltiples dimensiones y variables (figura 1.1), su selección y su priorización, en un proceso que mezcla planteamientos desde lo epistemológico, lo ideológico, hasta lo técnico. Por lo tanto, realizar una evolución implica previamente planificar dicha acción y elaborar el diseño permite en relación

con la realidad a evaluar. Por ello, se torna necesario reparar en las dimensiones afectadas.

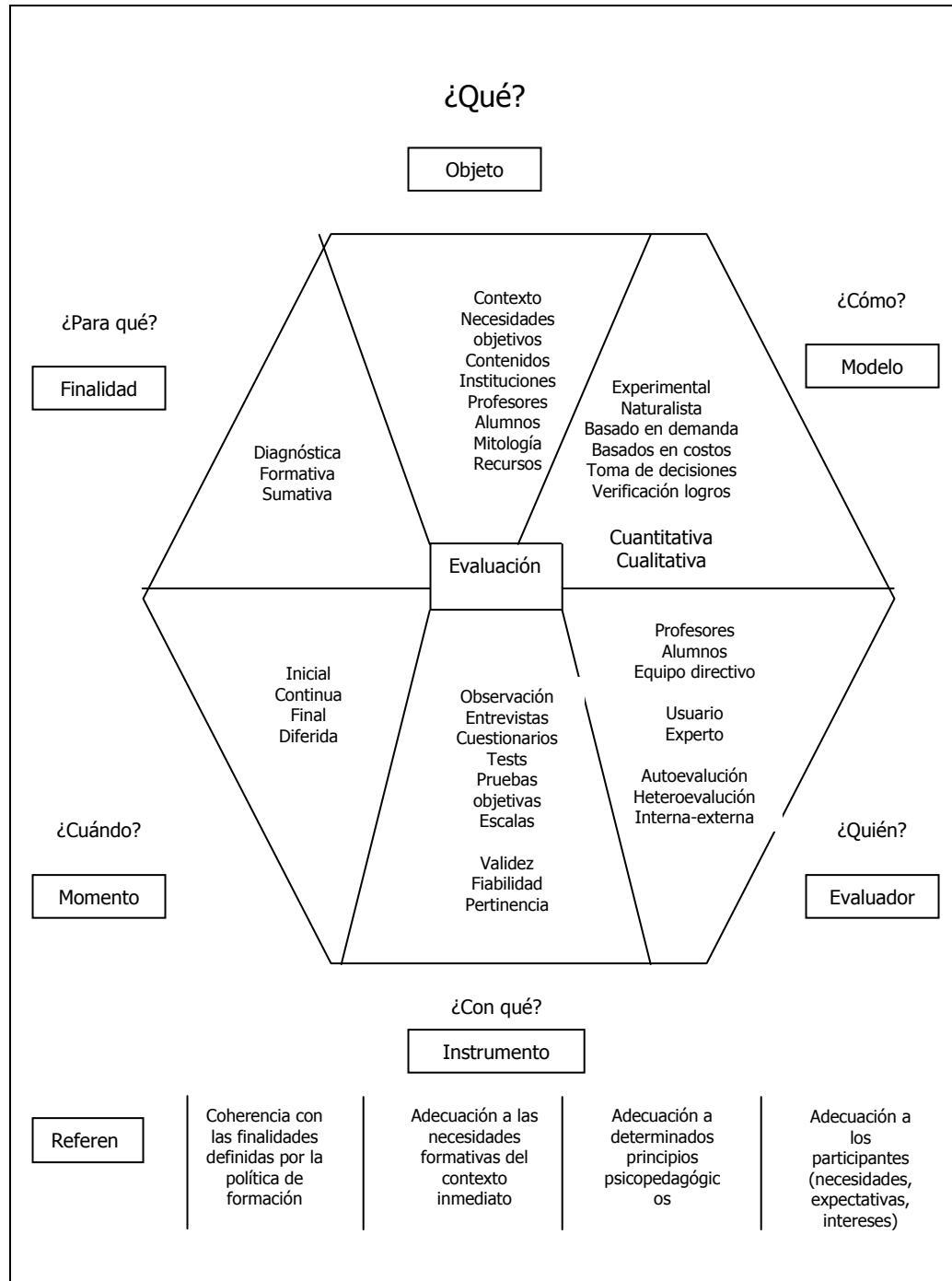


Figura 1.1. Dimensiones básicas de la evolución educativa (Tejada, 1991: 88).

Un primer acercamiento a las dimensiones afectadas en la evolución educativa se concreta en el diagrama de la figura 1.1. Un comentario sucinto del mismo nos indica que:

- a) *El objeto de evolución* puede ser diverso, aunque la mayoría de autores llegan a coincidir en torno a los alumnos, el personal docente y no docente, el currículum, los recursos. Los planes y programas, los proyectos, las estrategias metodológicas, las instituciones de formación y los contenidos entre otros con carácter general. Con posterioridad nos centramos, desde la óptica de ablución de programas, en algunos de los citados, para cada uno de ellos las fuentes de información, los protagonistas y estrategias de evaluación, obviamente, pueden variar en su nominación, implica y diseño.
- b) *La finalidad* perseguida por la evolución (diagnóstica, formativa y sumativa), aunque igualmente se podría hacer referencia a funciones más específicas (De Ketele y Roegiers, 1995) relativas a las personas (su rendimiento. Inicial, procesual o final), relativas al sistema (gestión), relativas al conocimiento (investigación), en la figura 1.2 se especifican las mismas y más allá de entrar en detalle en cada una de ellas conviene, en este momento, centrarnos en la evolución diagnóstica, la formativa y sumativa, como aglutinadoras de ellas.

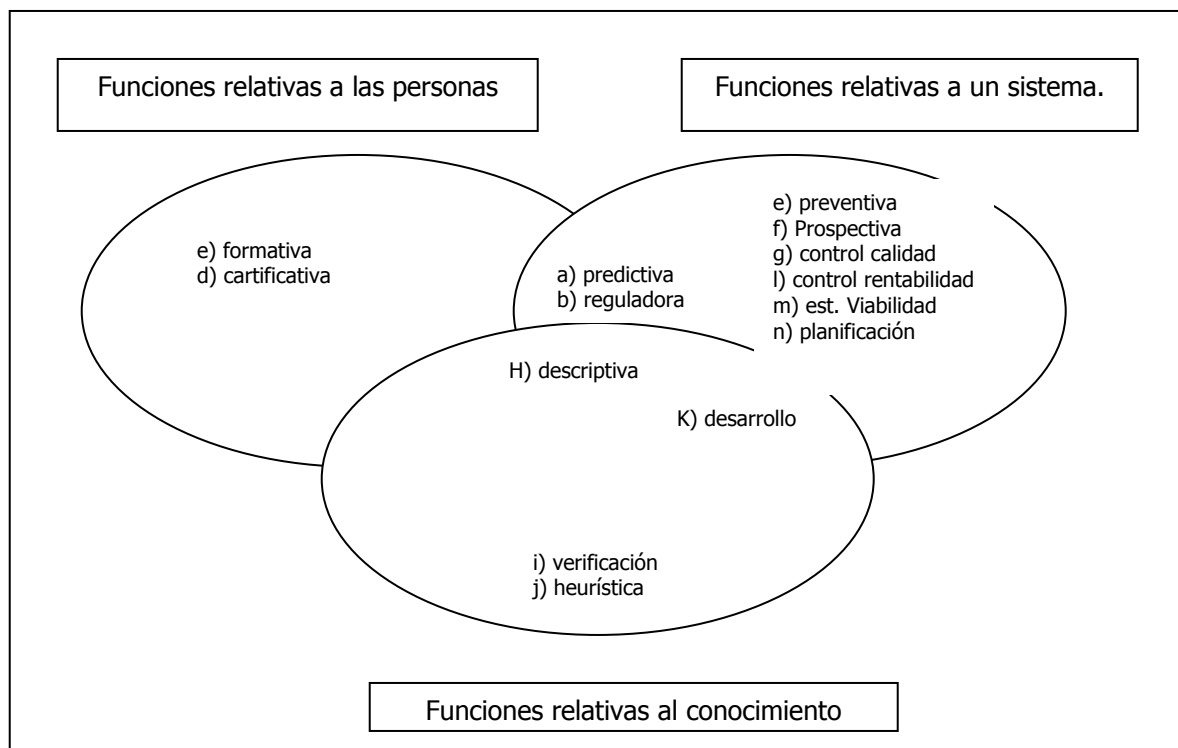


Figura 1.2. Las funciones de la evaluación (De Ketele y Roegiers, 1995: 134).

La evolución *diagnóstica* nos permite conocer la realidad en la que opera un programa (evolución de entrada), tanto en lo relativo al contexto como a los destinatarios del programa, ello nos da pie tanto a posibilitar los recursos y medios necesarios en el contexto para que el programa sea exitoso o bien, teniendo presente las limitaciones contextuales, ajustar el programa a dicho contexto. Por lo que se refiere a los destinatarios nos permite conocer sus necesidades formativas, sus posibilidades para con el programa, etc., lo cual facilita la toma de decisiones sobre admisión, orientación, establecimiento de grupos de aprendizaje según niveles, etc.

La evaluación *formativa* tiene que ver con el desarrollo del programa (evolución de proceso) en un contexto determinado y si propósito está directamente relacionado con la mejora y optimización del mismo en el transcurso de la acción formativa. Permite, pues, el ajuste racional de dicha acción según las vicisitudes que la misma encuentra y afecta básicamente a las estrategias de desarrollo (implementación de nuevos medios y recursos, ajuste de la temporalización – acelerando o lentificando el ritmo previsto-, cambio de actividades y papel del formador, etc.).

La evolución *sumativa* (evolución de productos) nos muestra los resultados obtenidos por el programa y en este sentido la toma de decisiones se orienta a la certificación, promoción, en relación con los destinatarios implicados en la acción formativa, a la vez que permite valorar globalmente el programa permitiendo rechazarlo o aceptarlo en función del éxito del mismo.

Expresamos en el cuadro 1.1. las relaciones e implicaciones entre estas dos últimas dimensiones.

- c) El *momento* evaluativo, conectado con la finalidad, es otra de las dimensiones básicas. Evaluación inicial con finalidad diagnóstica, al principio de una acción formativa, evolución continua con finalidad formativa, durante el periodo de formación y evaluación final con finalidad sumativa, al final del periodo de formación. Habría que añadir a estos tres momentos, ya clásicos, la evaluación diferida que en la evaluación de programas toma capital importancia por las repercusiones prácticas y sociales de la misma. Un programas no sólo es bueno por cuanto se verifica la suficiencia de los participantes al final de la acción formativa, sino también por cuanto a lo largo del tiempo tienen una repercusión en la práctica, también denominada *evolución de impactó del programa*.

Téngase presente que no siempre se evidencia el paralelismo establecido, por cuanto podemos verificar en la práctica que muchas veces se realiza evaluación continua con finalidad sumativa, en la medida que se recaba información no tanto para optimizar el programa sino más bien sobre el alumno para poderse asegurar

mejor en su evolución final, es decir, se van registrando resultados parciales más que la evolución del proceso formativo mismo.

Finalidad	Momento	Objetivos	Decisiones a tomar
Diagnóstica	Inicial	<ul style="list-style-type: none"> - Identificar las características de los participantes (intereses, necesidades, expectativas). - Identificar las características del contexto (posibilidades, limitaciones, necesidades, etc.). - Valorar la pertinencia, adecuación y viabilidad del programa. 	<ul style="list-style-type: none"> - Admisión, orientación, Establecimiento de grupos de aprendizaje. - Adaptación-ajuste e implementación del programa.
Formativa	Continua	<ul style="list-style-type: none"> - Mejorar las posibilidades personales de los participantes. - Dar información sobre su evolución y progreso. - Identificar los puntos críticos en el desarrollo del programa. - Optimizar el programa en su desarrollo. 	<ul style="list-style-type: none"> - Adaptación de las actividades de enseñanza-aprendizaje (tiempos, recursos, motivación, estrategias, rol docente, etc.).
Sumativa	Final	<ul style="list-style-type: none"> - Valorar la consecución de los objetivos así como los cambios producidos, previstos o no. - Verificar la valía de un programa de cara a satisfacer las necesidades previstas. 	<ul style="list-style-type: none"> - Promoción, certificación, reconsideración de los participantes. - A aceptación o rechazo del programa.

Cuadro 1.1. Relaciones entre finalidad y momento evolutivo.

d) El *modelo paradigmático*, que afectará al diseño a seguir. Éste está en íntima conexión con el referente (respecto a qué evaluamos). Condiciona la interacción entre las viables implicadas. Así, por ejemplo, los criterios para juzgar el valor o mérito de un objeto de evolución pueden ser la confluencia ente los hechos- logros- y los objetivos- modelos de evolución por objetivos-, la identificación de las necesidades actuales y potenciales de las audiencias, los ideales o valores sociales el estándar de conocimiento estructurados por los expertos u otros grupos relevantes. El cuadro 1.2 muestra una breve síntesis de los modelos más conocidos en evolución de programas.

Además, el modelo de interpretación puede basarse en la racionalidad positivista (científico-técnico), en la asignación de valores de la realidad

(interpretativo-fenomenológico) o enfatizar el sentido político que adopta cualquier situación (socio-crítico).

- e) La instrumentalización nos remite a las herramientas que han de servir para obtener la información considerada como necesaria para la evaluación. Instrumentos y técnicas de recogida de información estructurada, semiestructurados o sin estructuras, susceptibles de tratamientos cuantitativos o cualitativos según el objeto, pero, en cualquier caso, válidos, fiables, pertinentes, etc., como características técnicas exigibles a cualquiera de ellos. En definitiva, complementariedad y triangulación de instrumentos.
- f) El *evaluador* se convierte en un eje igualmente importante en el hecho evaluativo. Además de los implicados en el proceso formativo, están los expertos, lo cual nos puede permitir diferenciar entre evolución interna y evolución externa, entre Heteroevaluación y Autoevaluación. Por la relevancia de la problemática generada en torno a esta dimensión, con posterioridad nuevamente repararemos sobre ella.

Al margen de las ventajas e inconvenientes de una y otra, suelen combinarse en la práctica y utilizarse como elementos de complementariedad. Con todo, lo que se evidencia cada día con más fuerza es la necesidad de capacitación técnica para poder evaluar. El conocimiento de técnicas, instrumentos, metodologías, etc., en aras de una pretendida "objetividad" (intersubjetividad) es un requisito formativo para los protagonistas de la evolución.

- g) *El referente* es la dimensión más diferenciadora de la evolución: fuente de criterios e indicadores de evolución, está conectada con el resto de las dimensiones. En la evolución de programa, no cabe duda de que el referente debe pasar por la coherencia con la política formativa, la contextualización, los principios psicopedagógicos del aprendizaje adulto y la consideración de los participantes como elementos nucleares dentro de los cuales existe toda una gama de indicadores para la evolución.

La complejidad de la evolución, pues, se hace evidente a la vista de lo analizados y justifica plenamente la afirmación de que una evaluación nunca es completa en su misma. Por este motivo, la planificación de la misma tiene que hacer énfasis en la determinación de los aspectos más significativos de cada realidad y momento evaluativo.

I. 5. Modelos de evaluación

El concepto de evolución ha sufrido grandes modificaciones, tal y como hemos verificado en los apartados precedentes, desde el primer intento sistemático de acotarla por parte de Tyler (1950). Su evolución nos muestra una múltiple

apertura, que como ya nos refería Pérez Gómez (1983b: 431), puede concretarse en cuatro aspectos:

- a) *Conceptual*, dando cabida en la evolución a resultados no previstos o acontecimientos imprevisibles.
- b) *De enfoque*, posibilitando tanto la recogida de datos sobre productos y procesos.
- c) *Metodología*, pasando de la rígida e inflexible estrategia formal para incluir procedimientos informales, llegando al pluralismo metodológico.
- d) *Ético-política*, ya que la evolución proporciona información a todos los implicados y recoge opiniones e interpretaciones de los mismos. De la evolución burocrática a la evolución democrática.

Cuadro 1.2. Modelos de evolución de programas (Tejada, 1989b: 114-116).

Modelo	Verificación de logros	Evolución responderte	Toma de decisiones	Sin referencia a metas	Evolución iluminativa	Crítica artística
Organizador previo	Objetivos	Problemas localizados	Situaciones decisivas	Necesidades y valores sociales	Problemas y necesidades y localizados	Efectos
Objetivo	- Relacionar los resultados con los objetivos	- Facilitar la comprensión de las actividades y su valoración en una teoría determinada y desde distintas perspectivas - Ayuda a los clientes.	- Proporcionar conocimientos suficientes y una base valorativa para tomar y justificar las decisiones.	-Juzgar los méritos relativos de bienes y servicios alternativos.	-Estudiar del programa, intentando clarificar cuestiones, ayudar al innovador y a otras partes interesadas a identificar los procedimientos y aspectos del programa que pueden conseguir los resultados deseados.	-Ofrecer una representación y valoración de la práctica formativa y sus consecuencias dentro de su contexto.
Audiencia	-Planificadores y directivos del programa.	- Comunidad y grupo de clientes en zonas locales y los expertos.	- Los que toman decisiones.	-Los consumidores.	- Innovadores y personas involucradas.	-Comunidad y grupo de clientes en zonas locales.
Cuestiones previas	-¿Qué alumnos alcanzan los objetivos?	- ¿Cuál es la historia de un programa, sus antecedentes, operaciones, efectos secundarios. Logros accidentales, resultados?	-¿Cómo planificar, ejecutar y reciclar para promover el crecimiento y desarrollo a un coste razonable?	- ¿Cuál es el valor de un programa, dados sus costes, las necesidades concretas de los consumidores y los valores de la sociedad en general?	- ¿Cómo opera o funciona un programa, cómo influye en él los contextos donde se aplica?	-¿Cuáles son las características y cualidades de los fenómenos, cuáles son las cualidades emergentes de una clase o programa que se está evaluando?

Cuadro 1.2. (Continuación).

Modelo	Verificación de logros	Evaluación respondente	Toma de decisiones	Sin referencias a metas	Evaluación iluminativa	crítica artística
Metodología	Cuantitativa - planificación -experimental -comparación entre grupos.	Cualitativa - Descripción - Estudio de casos.	Cuantitativa - Planificación experimental - Valoración de necesidades - Observación - Estudios piloto.	Cuantitativo-cualitativa. - Planificación experimental. - Evaluación sin metas. - Análisis de coste. - Comparación-experimentación.	Cualitativa -Estudios de casos.	Cualitativa -Estudio de casos.
Instrumentos y técnicas	-Pruebas objetivas -Tests estandarizados de rendimiento pedagógico. -Observación predeterminada.	- Informes - Entrevistas. - Sociodrama. - Observación.	- Revisión de documentos. -Audiciones. - Entrevistas - Tests diagnósticos - Visitas - Escala de autovaloración.	- Listas de control. - Test diagnóstico. -Entrevistas. - Informes, etc.	- Observación - Entrevista	- Crítica artística.
Pioneros	Tyler	Stake	Stufflebeam	Scriven	Partett y Hamilton	Eisner
Proceso	- Establecer metas u objetivos. - Definir los objetivos en términos de comportamiento. - Establecer o desarrollar técnicas e instrumentos. - Recopilar datos. - Comparar los datos con los objetivos.	- Descripción antecedentes transacciones resultados. - Juicios antecedentes transacciones resultados.	-Análisis de tareas. - Plan de obtención de información. - Plan para el informe sobre resultados. - Plan para la administración del estudio.	- Valoración de necesidades. - Evaluación de metas. - Comparación con otras alternativas. - Examinar según coste y efectividad, -Combinación del personal con la del programa.	-Observación. -Investigación - Explicación.	-Descripción - Interpretación -Valoración.

Todo ello obedece al surgimiento de los distintos modelos de evaluación que podemos clasificar, pues según paradigmas enfoques, metodologías, etc. Distintos autores se han preocupado de este quehacer y han intentado sintetizar sobre la proliferación existente sobre el particular, ofreciendo sus propias v clasificaciones (Stake, 1976; Popham, 1980; House, 1980 y 1994; Davis, 1981; Guba y Lincoln, 1982 a; Nuevo, 1983; Srufflebeam y Shinkfield, 1987; Pérez Gómez, 1983; González y Escudero, 1987; Tejada, 1989 a y b; Garanto, 1989; Rosales, 1990; Avira, 1921; Zabalza, 1990).

Nosotros no vamos a analizar todos los modelos contemplados en las diferentes clasificaciones, sino que nos centraremos en los más representativos.

Previamente repararemos en torno a evolución cuantitativa y cualitativa dada su transferencia y relevancia en relación con los diferentes modelos.

I.5.I. Evolución cuantitativa-evaluación cualitativa

Básicamente dos son los grandes bloques de métodos en los que podríamos centrar la atención: los métodos cuantitativos y los métodos cualitativos. Los primeros históricamente predominantes hasta la década de los sesenta, y los segundos, que surgen como alternativos por la insuficiencia e indeterminación de los primeros para responder a las expectativas de los administrados, gestores y realizadores de programas de formación, aparecen porque los anteriores no logran encontrar suficientes respuestas sobre el funcionamiento, proceso, resultados e impacto de la formación, ya que los métodos cuantitativos se centran en la evolución de logros de los objetivos, desde una concepción restrictiva de la evaluación, referida simplemente al éxito en términos de adquisiciones observables, sinónimos de rendimiento académico.

La diferencia entre ambas líneas metodológicas no es solamente el giro realizado en torno al objeto de evolución proceso de recogida y tratamiento de la información, técnica de investigación, sino que es más profunda. Cada uno de ellos entraña un discurso epistemológico diferente, se fundamenta en paradigmas distintos y antagónicos, que implican distintas concepciones de la realidad, diversas maneras de en tender la ciencia o la naturaleza del conocimiento científico e incluso de los procedimientos para comprender los fenómenos humanos.

Todo ello ha generado una polémica entre evolución cuantitativa y evolución cualitativa, que en el terreno que nos ocupa-evaluación en el ámbito no formal- se agudiza y cobra más sentido. En la década de los ochenta ha tenido bastante profusión y se le ha dedicado muchas páginas a la polémica p'planteada. Desde posturas irreconciliable entre paradigmas, distintos trabajos (Parlett y Hamilton, 1977; Stake, 1986; Guba, 1983; etc.) defienden los métodos cualitativos, mientras que otros (Cronbach y otros, 1980; Riecken y otros, 1974; Popham, 1980, etc.) fundamentarían los métodos cuantitativos, existiendo posturas intermedias o de

reconciliación (Cook y Reichardt, 1986; Patton, 1984; Smith y Heshusius, 1986, etc.).

Lo que sí parece cierto es que dada la complejidad que presentan los fenómenos educativos es necesario recurrir a la triangulación de metodologías definida como la combinación de las mismas.

No insistimos más en este tipo de argumentaciones en torno a la complementariedad e integración entre lo cuantitativo y lo cualitativo que tiende a prevalecer y que personalmente asumimos, con independencia de los problemas de orden técnico, tiempo, recursos, capacitación y formación del evaluador, etc. El en cuadro 1.3 nos limitaremos a exponer las características más sobresalientes de cada una de estas posturas evaluativas, con el objeto de arrojar luz suficiente, para más tarde poder ubicar los distintos modelos de evolución que se han generado a raíz de tales premisas.

I. 5.2. Diferentes modelos de evaluación

A) Modelo de evolución orientada hacia los objetivos

Se puede considerar *la evolución como el primer modelo de evaluación sistemática. Tiene su origen en los treinta y a R.W. Tyler como su representante principal.*

Se caracteriza la evaluación como el proceso de determinar de determinar la congruencia entre los objetivos y los logros. Tyler (1950: 69) afirma que "el proceso de evolución es esencialmente el proceso de determinar hasta qué punto los objetivos educativos han sido actualmente alcanzados mediante los programas de currículum y enseñanza".

Este planteamiento, basado en la congruencia, parte de la base lógica organizada a partir del desarrollo de un programa, dentro de una concepción tecnológica de la enseñanza. Es más, el propio Tyler consideraba la evaluación como una etapa esencial en el desarrollo., además consideraba dentro de esta etapa una serie de pasos que determinan el proceso real de evaluación.

Cuadro 1.3. Comparación entre evolución cuantitativa (tejada, 1989b: 136-137).

Evolución cuantitativa	Evolución cualitativa
1. Naturaleza de la realidad	
- Estable y fina. - Unidad en la naturaleza, Mecánica y unidireccional	- Dinámica y cambiante - Rechaza la unidad de la naturaleza
2. Relación objeto-sujeto	
- El individuo como sujeto pasivo de la realidad. - Interesada por la búsqueda de hechos o causas, centrándose en las conductas observables- positivismo lógico-.	- El individuo como agente constructor de la realidad en la que está inmerso. - Interesado por comprender no sólo las conductas manifiestas (procesos de pensamiento, interpretación y significación de las personas implicadas en la evolución). - Fenomenologismo.
3. Finalidad de la evaluación	
- La evolución se refiere al éxito o fracaso del logro de los objetivos. - Énfasis en los productos.	- La evolución no sólo se refiere a objetivos, sino a necesidades y valores sociales, solución de problemas localizados. - Énfasis en los procesos,
4. Contexto	
- La evolución es ajena al contexto.	- La evolución debe tener en cuenta las particularidades de cada contexto.
5. Papel del evaluador	
- El evaluador está al margen de los datos con una perspectiva teórica "desde fuera".	- El evaluador necesita esquemas explicativos para acercarse a cada realidad. Su perspectiva en relación a los datos es "desde dentro" lo que le permitirá no sólo comprender sino también elaborar una explicación de los fenómenos en relación con su ocurrencia en la realidad.
6. Metodología	
- Se basa en el método hipotético deductivo, comparación de grupos, comprobación, etc.	- Se basa en el método inductivo, exploratorio, descriptivo, expansivo, estudio de casos, etc. - La estrategia es holística.
7. Diseño	
- La estrategia es particularista. - Diseños y planes rígidos.	- Diseño flexibles y se configuran con los datos.
8. Instrumentos	
- Instrumentos básicos son tests estandarizados y las pruebas objetivas, observación sistemática.	- Los instrumentos son múltiples (entrevista, cuestionarios, etnografías, informes, trabajos, etc.).

En este sentido, el procedimiento del autor para evaluar queda reflejado en los pasos siguientes:

1. Establecer las metas u objetivos.
2. Ordenar los objetivos en amplias clasificaciones.
3. Definir los objetivos en términos de comportamiento.
4. Establecer situaciones y condiciones según las cuales puede ser demostrada la consecución de los objetivos.
5. Explicar los propósitos de la estrategia al personal más importante en las situaciones más adecuadas.
6. Escoger o desarrollar las apropiadas medidas técnicas.
7. Recopilar los datos de trabajo.
8. Comparar los datos con los objetivos de comportamiento.

El modelo exige, como puede observarse, la operativización de los objetivos, siendo éstos, además, la fuente de criterio de evolución. Por otra parte, los resultados educativos se concretan en términos de criterios específicos de ejecución del alumno, que se mide por medio de test estandarizados, los cuales se analizan estadísticamente comparando todos los alumnos, siendo los criterios de logros los mismos para todos, los alumnos (Patton, 1984: 51). En consecuencia, la elaboración de baterías de tests o pruebas de rendimiento para examinar la actuación del alumno antes después del programa constituye también una tarea importante dentro del proceso evaluativo. De todo lo anterior se desprende una información útil para el equipo que se encarga de diseñar el programa o para aquellos que se interesen por su aplicación.

Las *ventajas* del modelo, dentro de la perspectiva teórica en la que se enmarca *_cuantitativa_*, se centran en que el modelo desplaza su atención a otros aspectos del programa, aparte de los propios alumnos (típico de la época pretyleriana) y en la toma de decisiones que permite una vez concluido el desarrollo del programa,. No cabe duda, que por su sistematismo ha sido el modelo más popular y utilizado dentro de la corriente de la pedagogía por objetivos. Sus limitaciones son varias y fruto de ello las consiguientes críticas:

- En primer lugar, hay que considerar la evaluación como un proceso terminal que proporciona información cuando se ha cubierto el ciclo del programa, con lo cual se pierde la posibilidad de rectificar el programa cuando se está desarrollando. La evolución formativa, pues, que es el aspecto más importante en el desarrollo de un programa, se sacrifica al rigor en términos de evolución sumativa.
- El énfasis en los objetivos preestablecidos provoca "el olvido e ignorancia intencional de los efectos secundarios de la escuela y de los acontecimientos accidentales, locales, no previstos. No siempre lo previsto

es lo más relevante y significativo en el curso de una experiencia educativa” (Pérez Gómez, 1983b: 434). Pero además, el evaluador debe seleccionar u clarificar los objetivos apropiados de manera operativa, con lo que se puede caer en la criba o selección de unos cuantos, por ello, los objetivos importantes, particularmente los relacionados con los aspectos menos concretos del programa, como son las actitudes que debían adoptarse, son con frecuencia olvidados (Stufflebeam y Shinkfield, 1987: 97). De hecho, se deriva cierta redenciosidad en esta fase de selección de objetivos de comportamiento.

- La consideración del rendimiento –logro- como último criterio implica que sea *un modelo reduccionista*, en el que se remplace el estudio de un comportamiento condicionado por el contexto, por la consideración de objetivos artificiales –impuestos desde fuera y con anterioridad-. Por tanto, no tienen presente que los logros del alumno sólo pueden comprenderse en un contexto concreto y en relación con los significados particulares que el propio alumno concede a su experiencia (Patton, 1984: 51).
- Se centra en *comportamientos observables* que no son sino indicadores simbólicos de procesos internos y aprendizaje complejos que no pueden deducirse directamente de manifestaciones externas.
- Es un *modelo burocrático* en el que la audiencia de la evolución es la productora del programa, no sus usuarios.

B) El modelo científico de evolución

Este modelo conlleva que la evolución de programad debe basarse en la lógica del método científico. Su máximo representante es Suchman (1967) que parte de la concepción de que el evaluador debe ser un investigador.

En este sentido, se entiende la evolución como:

[... un proceso social continuo, que supone intrínsecamente una combinación de suposiciones básicas acerca de la actividad que está siendo evaluada y los valores personales de quienes participan en el estudio, incluyendo el, evaluador, insistiendo en que la evolución debe s necesariamente convertirse en un proceso científico que dé cuenta de esta subjetividad, ya que no puede ser eliminada (Strufflebeam y Shinkfield, 1987: 112)].

Los principios o propósitos evaluativos de este modelo son.

1. Describir su los objetivos han sido alcanzados y de qué manera.
2. Determinar las razones de cada uno de los éxitos y fracasos.
3. Describir los principios que subyacen en el programa que ha tenido éxito.
4. Dirigir el curso de los experimentos mediante técnicas que aumenten su efectividad.

5. Sentar las bases de una futura investigación sobre las razones del relativo éxito de las técnicas alternativas.
6. Redefinir los medios que hay que utilizar para alcanzar los objetivos, así como las submetas, a la luz de los descubrimientos de la investigación.

En estos propósitos puede observarse una estrecha relación entre la planificación y desarrollo del programa y su evolución. La investigación proporcionará la información básica para planificar y, si es necesario, replanificar. No obstante, Suchman también admite que en la planificación y realización de un programa, muchas de las cuestiones de evolución pueden ser resueltas sin necesidad de investigación. De todas formas, siempre que los requisitos de la investigación evaluativa puedan ser satisfechos, los resultados serán conclusiones basadas en investigaciones convincentes y no en juicios subjetivos.

Los criterios de éxito o fracaso para evaluar un programa, según Suchman, pueden ser:

- *Esfuerzo*. Las evoluciones de este tipo presentan la cantidad y la calidad de las actividades producidas durante el desarrollo del programa. La valoración implica lo que se hace y no el producto.
- *Trabajo*. Se valoran más los resultados del esfuerzo en sí mismo.
- *Suficiencia* del trabajo. Considerando como el grado en que el trabajo realizado coincide, por comparación, con todas las necesidades de acuerdo con los objetivos.
- *Eficiencia*. Relación entre el esfuerzo realizado y los resultados.
- *Proceso*. Investiga las explicaciones básicas las razones que conducen a los resultados. Las dimensiones para el análisis del proceso son: a) los atributos del programa, b) la ente a la que afecta el programa, c) el contexto en el que se desarrolla el programa, y d) los diferentes tipos de efectos producidos por el programa.

Las ventajas reconocidas al modelo se centran en la propia metodología de evolución, al asumir la misma como un proceso son continuo y abierto, que se aborda desde una actitud crítica y no dogmática o preestablecida. No debemos olvidar en este sentido, sin dejar su referente teórico dentro del planteamiento tecnológico, que se trata de recoger y analizar datos para demostrar el valor de la actividad social, y propiciar con ello decisiones administrativas más objetivas. Sus *limitaciones* igualmente son varias y pueden centrarse en torno a:

- La utilidad *de los resultados*, muchas veces está motivada por "intereses creados", y sobre todo por los propios administradores o gestores de programas educativos.
- El destinatario *del programa* se sigue considerando más desde un planteamiento conductual y así se siguen entendiendo los objetivos, de

manera que sus actitudes, motivaciones, etc., no se tienen en cuenta para calibrar la valía del programa.

- El evaluador con sus percepciones, creencias, o prejuicios pueden mediar en el análisis de los resultados.
- La instrumentalización y el proceso evaluativo quedan configurados desde un planteamiento eminentemente cuantitativo e incluso en situaciones experimentales en aras a la pretendida validez y fiabilidad de todo el proceso de investigación.

C) El modelo de toma de decisiones

En este modelo la evolución se estructura y organiza en relación con las decisiones que se han de tomar a lo largo del desarrollo de un programa.

Su principal representante es Stufflebeam, cuya definición se recogió en el epígrafe 1.4.

En esta definición se resumen los conceptos clave del modelo CIPP (Context-Input-Process-Product), cuyo proceso de evolución rearticula básicamente en tres pasos:

1. *Delimitar*: los evaluadores se reúnen con las personas responsables de la toma de decisión en orden a identificar la información necesaria.
2. *Obtener*: recoger y procesar la información.
3. *Aplicar*: suministro de la información recogida y procesada a los responsables de las decisiones para que ésta sean más racionales.

Stufflebeam distingue cuatro tipos de decisiones en estrecha relación con los tipos de evolución que nos plantea: contexto, entrada, proceso y producto. En el cuadro 1.4 expresamos dicha relación.

Para resumir las ventajas de este modelo, podemos decir que resuelve el problema del evaluar porque toma como destinatario (a quien se dirige la evolución) al responsable de las decisiones y adopta como criterios significativos los propios de este último. Las cuestiones que se resuelven corresponden a los responsables de las decisiones, aunque, por regla general, incluyen la eficacia del programa en alguna dimensión y, en especial, qué partes del programa funciona.

Por el contrario, el olvido del carácter jerárquico de los sistemas educativos y la tendencia a la imposición desde arriba y desde afuera se convierten en los puntos críticos y por tanto en sus principales *limitaciones*.

D) Modelo de evolución responderte

Este modelo elaborado por Stake (1976) pone el énfasis en la acomodación a las necesidades de los "clientes". Parte de la idea de que las intenciones pueden

cambiar y pedía una comunicación continua entre el evaluar y la audiencia con el fin de descubrir, investigar y solucionar los problemas. De manera que es necesario describir ofrecer un retrato completo y holístico del programa educativo. Las características principales de esta concepción son.

- Las evoluciones deben ayudar a las audiencias a observar y mejorar lo que están haciendo.
- Los evaluadores deben describir programas con relación tanto a los antecedentes y las operaciones como a los resultados.
- Los efectos secundarios y los logros accidentales deben ser tan estudiados como, los resultados buscados.
- Toma en consideración las diferentes interpretaciones de aquellas personas que están implicadas en el programa.
- Los evaluadores deben evitar la presentación de conclusiones finales resumidas, pero en su lugar deben recopilar, analizar y reflejar los juicios de una amplia gama de gente interesada en el objeto de la evolución.
- Los experimentos y los tests estandarizados son a menudo inadecuados o insuficientes para satisfacer los propósitos de una evolución, y deben ser frecuentemente sustituidos o completados con una variedad de métodos.

La estructura funcional de la evolución respondete se concreta e 12 pasos, teniendo en cuenta que no es una estructura lineal:

1. Hablar con los clientes, el personal del programa y las audiencias.
2. Identificar el alcance del programa.
3. Panorama de las actividades del programa.
4. descubrir los propósitos e intereses.
5. conceptuar las cuestiones y problemas.
6. Identificar los datos necesarios para investigar los problemas.
7. Seleccionar observadores, jueces e instrumentos si los hay.
8. Observar los antecedentes, las transacciones y los resultados propuestos.
9. desarrollar temas, preparar descripciones y estudiar casos concretos.
10. Validación: confirmación, búsqueda de evidencias para la no confirmación.
11. Esquema para el uso de la audiencia.
12. Reunir los informes formales, si los hay.

Cuadro 1.4. Cuadro de evolución (Stufflebeam, 1987: 194-195).

	Contexto	Entrada	Proceso	Producto
Objetivo	Definir el contexto institucional, identificar la población objeto de estudio y valorar sus necesidades. Identificar las necesidades, diagnosticar los problemas que subyacen en las necesidades y juzgar si los objetivos propuestos son suficientemente coherentes con las necesidades valoradas.	Identificar, y valorar la capacidad del sistema, las estrategias del programa, las alternativas, la planificación de procedimientos para llevar a cabo las estrategias, los presupuestos y los programas.	Identificar o pronosticar, durante el proceso, los defectos de la planificación del procedimiento o de su realización, proporcionar información para las decisiones preprogramadas y describir y juzgar las actividades y aspectos del procedimiento.	Recopilar descripciones y juicios acerca de los resultados y relacionarlos con los objetivos y la información proporcionada por el contexto, por la entrada de datos y por el proceso e interpretar su valor y su mérito.
Método	Utilización de métodos como el análisis de sistemas, la revisión de documentos, las audiciones, las entrevistas, los tests diagnósticos, la técnica, Delphi...	Inventariar y analizar los recursos humanos y materiales disponibles, las estrategias de solución y las estrategias de procedimiento referentes a su aplicabilidad, viabilidad y economía. Y utilizar métodos como la búsqueda de bibliografía, las visitas a programas ejemplares, los grupos de asesores y ensayos piloto.	Controlar las limitaciones potenciales del procedimiento y permanecer alerta ante las que nos e esperaban mediante la obtención de información específica de las decisiones programadas, la descripción del proceso real, la continua interacción con el personal del proyecto y la observación de sus actividades.	Definir operacionalmente y valorar los criterios de los resultados, mediante la recopilación de los juicios de los clientes y la realización de análisis cuantitativos y cualitativos.
Relación con la toma de decisiones	Decidir el marco que debe ser abarcado, las metas relacionadas con la satisfacción de las necesidades o la utilización de las oportunidades y los objetivos relacionados con la solución de problemas. Y proporcionar una base para jugar los resultados.	Seleccionar los recursos de apoyo, las estrategias de solución y las planificaciones de procedimientos, esto es, estructurar las actividades de cambio. Y proporcionar una base para juzgar la realización.	Llevar a cabo y perfeccionar la planificación y los procedimientos del programa, esto es, efectuar un control del proceso. U proporcionar un esbozo del proceso real para utilizarlo más tarde en la interpretación de los resultados.	Decidir la continuación finalización, modificación o readaptación de la actividad del cambio. Y presentar un informe claro de los efectos (deseados y no deseados, positivos y negativos).

Las *ventajas* del modelo responderte se pueden vislumbrar a partir de que el mismo ha generado toa una escuela de evolución que exige un método pluralista, flexible, holístico, subjetivo y orientado hacia el cliente, en donde el evaluador es un "buscador de la verdad", asumiendo que existen múltiples, y que los evaluadores pueden elegir distintos caminos para revelar la utilidad del programa. Stake cree que la evolución interna es más útil que la externa y la Autoevaluación mejor que la evolución centralizada o heteroevaluación.

Igualmente cabría señalar algunas limitaciones al modelo. En este sentido destacamos la escasa consideración en la precisión de los datos, la dificultad de distinguir entre descripción y análisis, la confianza en la capacidad de los participantes para analizar las interacciones complejas y los factores latentes y la tendencia a infravalorar la naturaleza y la función política de la evolución.

E) Modelo de evolución iluminativa

Este modelo, propuesto por Parlett y Hamilton (1977) surge como alternativo a los modelos convencionales de corte cualitativo. Las características más relevantes del modelo son (Pérez Gómez, 1983b: 441):

- Los estudios sobre evolución deben comprender una tendencia holística y tener en cuenta el amplio contexto en el que funciona.
- Se preocupan más de la descripción e interpretación que de la medida y la predicción.
- Se orienta al análisis de los procesos más que al análisis de los productos.
- La evolución se desarrolla bajo condiciones naturales o de ampo y no bajo condiciones experimentales.
- Los métodos principales de recogida de datos son la observación y la entrevista.

Las metas de la evolución iluminativa se pueden concretar en.

1. *Estudiar el programa* teniendo en cuenta cómo opera o cómo funciona, cómo influyen en él las situaciones educativas en las que se aplica, ventajas u desventajas, cómo se ven afectadas las tareas intelectuales y las experiencias académicas de los participantes.
2. *Descubrir y documentar* qué significa participar en el programa, bien sea como formador o como participante.
3. *Discernir y comentar* las características más significativas del programa, las concomitancias recurrentes y los procesos críticos.

De esta forma "se intenta plantear y clarificar una compleja serie de cuestiones, ayudando al innovador o a otras partes interesadas a identificar aquellos

procedimientos y aspectos del programa que pueden conseguir los resultados deseados” (Stufflebeam y Shinkfield, 1987: 320).

En este modelo aparecen dos conceptos de capital importancia: el *contexto de aprendizaje* y el *sistema de instrucción*. El primero hace referencia al contexto material, psicológico y social dentro del cual trabajan conjuntamente formadores y participantes; representan una red de variables culturales, sociales, institucionales y psicológicas que interactúan de forma compleja. El segundo concepto, sistema de instrucción, como producto teórico es un conjunto de elementos que componen un plan coherente, que al aplicarse sufre modificaciones en función del medio, los participantes y los formadores; por ello deben evaluarse en función de las condiciones concretas de su aplicación.

La evolución iluminativa no es un método regularizado sino una estrategia global que puede adoptar diversas formas. Pretende ser adaptable para descubrir las dimensiones, metas y técnicas de evolución, y ecléctica en el sentido de proporcionar al evaluador una serie de tácticas investigativas, que dependerán del problema que deba investigarse. La preocupación del evaluador será familiarizarse con la realidad cotidiana en la cual está investigando, sin intentar manipular, controlar o eliminar las variables situacionales, si no abarcan toda la complejidad del contexto.

Dado que las evoluciones iluminativas subrayan el examen del programa como parte integrante del contexto de aprendizaje, metodológicamente se hace hincapié en la observación de las aulas y las entrevistas con los profesores y alumnos participantes, así como los cuestionarios y el análisis de documentos. El diseño no se establece de antemano, sino que al ser flexible se va acomodando y puede modificarse con el curso de la acción.

Parlett y Hamilton (1977) distinguen tres fases en la evolución iluminativa:

1. *Observación*: se investiga toda una amplia gama de variables que afectan al resultado del programa.
2. *Investigación*. El interés de desplaza desde el reconocimiento hacia la selección y el planteamiento de cuestiones de manera coherente, con el fin de realizar una lista sistemática y selectiva de los aspectos más importantes del programa en su contexto.
3. *Explicación*: en la que los principios generales subyacentes a la organización son expuestos y se delimitan los modelos causa-efecto e sus operaciones.

Una de las *ventajas* de este modelo es su versatilidad, des decir, las diferencias entre los informes es una inevitable y deseable consecuencia de este método evaluativo.

Igualmente la evolución resalta más la recopilación de información que la toma de decisiones. La tarea de quien prepara el informe, para que se tomen las decisiones, es proporcionar un amplio aspecto de complejas realidades a cerca del programa. De ahí deben surgir discusiones que conduzcan a desentrañar dichas complejidades y aislar lo trivial de lo significativo.

Sin embargo, también se le vislumbran algunas *limitaciones*. Stenhouse (1984) critica a los evaluadores iluminativos porque los criterios que empelan para verificar el calor o el mérito del currículum o la práctica educativa no son claros y su preocupación por las audiencias y la presentación de los resultados no son más que un enmascaramiento del problema. Por otra parte, la utilización de técnicas abiertas, datos cualitativos, etc., sigue permitiendo una gran parcialidad por parte del investigador, a la par que puede incidir en la realidad en la que opera "desde afuera" más que "desde dentro".

F) Modelo de evolución sin referencia a los objetivos (metas)

Este modelo de evolución, propuesto por Estriben (1973), tienen la peculiaridad en principio, de obviar los objetivos del programa, creyendo que con esa ignorancia por parte del evaluador, será más objetivo al evaluar los resultados del programa, tanto los previstos como los no previstos.

El propio autor parte de la idea de evolución como una actividad metodológica que "consiste simplemente en la recopilación y combinación de datos de trabajo mediante la definición de unas metas que proporcionen escalas comparativas o numéricas, con el fin de justificar los instrumentos de recopilación de datos, las valoraciones y la sección de metas" (Soriven, 1967). Quiere ello decir que la evolución es la determinación sistemática y objetiva del valor o el mérito de un objeto basándose en pruebas acumuladas procedentes de su comparación con otros objetos distintos a la hora de satisfacer las necesidades del consumidor. El modelo de Scriven se desarrolla de acuerdo a lo siguiente:

1. Antecedentes, Contexto y, recursos y función del programa/ servicio que se quiere evaluar.
2. El sistema de distribución del programa/intervención.
3. Descripción de los usuarios o población objeto de intervención.
4. necesidades y valores de los que se ven realmente afectados por la intervención
5. Existencia o no de normas/criterios previos para la evolución.
6. El proceso de intervención social.
7. Los resultados de la misma.
8. los costes de la intervención.

9. La comparación con programas/servicios alternativos.

No cabe duda de que en este modelo se rompe la tradición tyleriana, puesto que para Scriven el modelo de objetivos es imperfecto y las evaluaciones propuestas potencialmente inútiles, ya que las metas y objetivos propuestos pueden ser inmortales, poco realista y no representativos de las necesidades de los consumidores o demasiado limitados como para prever los efectos secundarios posiblemente cruciales. En este sentido. El autor afirma que los evaluadores deben juzgar las mentes, pero en limitarse a ellas a la hora de buscar los resultados.

De cualquier forma, insiste también Scriven, la evolución sin metas es reversible y complementaria, es decir, se puede empezar sin metas con el fin de investigar todos los efectos, y luego cambiar de método para asegurar que la evolución determinará si las metas han sido alcanzadas, o incluso pueden realizarse simultáneamente ambos tipos de evolución por diferentes evaluadores.

Para Scriven existen dos funciones capitales de la evolución. *La formativa*, que ayuda a desarrollar el programa, *la sumativa*, que calcula el valor del objeto una vez que ha sido desarrollado. La evolución formativa es parte del proceso de desarrollo, proporciona información continua para ayudar a planificar de cara a perfeccionar cualquier aspecto que esté operado. La evolución sumativa puede servir para ayudar a los administradores o consumidores a decir si todo el currículum, programa, etc., ya finalizado (mejorado a través de la evolución formativa) representa igualmente un avance sobre otras alternativas disponibles lo suficientemente significativo como para justificar los costes de su adopción.

De todo lo anterior se desprende una serie de ventajas que su autor subraya. es una evaluación menos intrusiva que la evolución basada en objetivos, más adaptable a los cambios de metas repentinos, más solvente a la hora de encontrar los efectos secundarios, menos propensa a la tendenciosidad social, perceptiva o cognitiva, más equitativa a la hora de tener en cuenta una amplia gama de valores, etc.

Pero entraña algunas dificultades que se convierten en sus limitaciones. La evolución sin metas conlleva un problema importante cual es la asignación de un significado valorativo a los resultados para verificar la eficacia y el valor de un programa. Además, Scriven introduce el concepto de necesidad como base de la evolución. Considerando que la misma se hará con referencia a las necesidades del consumidor más que a los objetivos e intenciones de su productor.

La valoración de las necesidades es un proceso racional, objetivo y externo, es decir, es ajeno a las preferencias subjetivas de las distintas audiencias. Por ello es discutible dentro del modelo de Scriven. Además, dados los resultados de una valoración de necesidades, pueden juzgarse los resultados de un programa concreto como buenos, malos o indiferentes, dependiendo de su contribuyen o no

a satisfacer las necesidades identificadas. Este argumento, afirma Stufflebeam y Shinkfield, "es más conveniente en el terreno filosófico que en el práctico. Se necesita un alto desarrollo de nivel técnico antes de que la valoración de necesidades pueda introducir significados factibles y apropiados para juzgar los resultados" (1987:348).

G) Modelo de evolución basada en la crítica artística

Este modelo de evolución, propuesto por E. Eisner, se fundamenta en una concepción de la enseñanza como un arte y el profesor como un artista,. Apoyándose en el currículum como una realidad cultural que encarna normas y reglas implícitas de la cultura.

La evolución del currículum se basa, por tanto, en la crítica, a partir de la interpretación y comprensión del contexto, de los símbolos, reglas y tradiciones de los que participan. La tarea del evaluar es desvelar la calidad de las situaciones curriculares cuando se transforman en intenciones y realidades. Los conceptos clave de este modelo son dos:

- *Connoisseurship*: capacidad de percepción cualitativa, considerando como "el arte de percepción que posibilita la apreciación de la complejidad de las prácticas educativas (Eisner, 1981). El evaluador, por su formación y experiencia, es capaz de apreciar y reconocer las características y cualidades de los fenómenos, lo que le permite ser sensible a las cualidades emergentes del programa que se está evaluando.
- *Crítica artística*: como "arte de revelación", el evaluador, además de discernir el carácter o cualidades de un acontecimiento, debe ofrecer una representación de lo que ha encontrado para que otras personas que no posean su nivel puedan participar de ello, comprendido y valorando la calidad de la práctica educativa y sus consecuencias.

La crítica artística se presenta en un documento escrito en el que pueden distinguirse tres aspectos (Eisner, 1981):

- Carácter descriptivo, ya que en el mismo se presenta una descripción viva y detallada de la situación evaluada desde una perspectiva fáctica (incluye referencias directas de las intervenciones de los participantes, profesores y alumnos) y artística (literario, metafórico).
- Carácter interpretativo, que conlleva la búsqueda de significados, un intento de entender lo que ocurre más allá de lo observable, de inferir y detectar interacciones complejas y latentes.

- *Carácter valorativo:* al emitir juicios de valor respecto de lo que se ha descrito e interpretando, teniendo presente que la valoración se realiza de acuerdo a las peculiaridades que singularizan y definen cada situación y no siguiendo criterios estándares o universales.

Como puede derivarse de todo lo anterior, este modelo, más allá de romper con la evolución tradicional, posee la ventaja de ser procesal y cualitativo al ser sensible a las cualidades que surgen en la vida del aula, de la institución educativa u otras realidades; tratando de capturar lo singular, lo relevante, las relaciones que definen el comportamiento de las personas y el desarrollo de los procesos.

Por el contrario se le asocian limitaciones sobre todo en relación con la finalidad y validez. Para Eisner, éstas se logran a través del proceso de "corroboración estructural" y "adecuación referencial". Es en ese punto donde se le encuentran debilidades al ser insuficientes sus procedimientos de justificación metodológica, de búsqueda de evidencias.

H) Modelo de evolución democrática

Este modelo, propuesto por MacDonald, se sitúa en un planteamiento eminentemente cualitativo, dentro de una visión naturalista de la realidad. Sólo abarcando la realidad como un todo vivo y en movimiento, tal y como se produce en el día a día, puede la investigación servir para la acción. Para conocer la realidad y sus significados, pues, hay que sumergirse en el curso real y vivo de los acontecimientos y conocer las interpretaciones diversas que hacen de la misma los que en ella participan. Esto implica que tanto las fuentes de datos como los destinatarios de los informantes son todos los que participan en un programa y, por tanto, el modelo de evaluación tiene que ser democrático (MacDonald, 1983). Las opiniones e interpretaciones de todos los afectados han de estar presentes u sólo un proceso de negociación que así lo considere es garante para llegar a la evaluación realista. Más concretamente podemos precisar algunas características.

- La evolución democrática es un servicio de información a la comunidad entera de las características de un programa educativo.
- El patrocinio de evolución no concede derecho especial sobre este servicio
- El evaluador democrático reconoce el pluralismo de valores y busca la representación de intereses diferentes en su formulación de cuestiones.
- El valor básico es la información a los ciudadanos, actuando el evaluador como agente de intercambio de información entre diferentes grupos.
- Las técnicas de recogida y presentación de datos deben ser accesibles a audiencias no especializadas (entrevistas, debates, etc.).
- La principal actividad del evaluador es la recogida de definiciones del programa y de reacciones frente a él.

- El informante no proporciona recomendaciones. Las interpretaciones y opiniones que contienen surgen de un proceso de negociación con todos los implicados en el estudio.
- Los conceptos clave de la evolución son "carácter confidencial", "negociación" y "accesibilidad". El concepto justificador es el "derecho a saber".

Algunas *limitaciones* al modelo, en consonancia con Pérez Gómez (1983b), las podemos encontrar en el poder y utilidad de la información. De manera que en una sociedad democrática, con realidades desiguales y jerarquizadas, ofrecer información a quien ostenta el poder supone directa o indirectamente fortalecer su posición. De ahí que la posición del evaluador difícilmente puede ser neutral ante las diversas audiencias. Por otra parte, puede resultar complejo conjugar la pluralidad u flexibilidad metodológica y procedimental y la transparencia expositiva a efectos informativos de las audiencias, con el necesario rigor y precisión de planteamientos y conceptos. Por último, hay que considerar el problema de la confidencialidad. Debe reconocerse que no resultará fácil lograr el equilibrio entre el derecho de la audiencia a conocer la realidad y el de los agentes participantes a mantener la intimidad de sus apreciaciones.

Actividades de desarrollo

- I. Intente redactar su propia definición personal, tras haber leído, pero sin revisar de nuevo los textos:
 - a) Trate de determinar el autor en el que más se ha apoyado.
 - b) Considerando los ocho modelos presentados, ¿con cuál encaja mejor conceptual y técnicamente?
- II. Escoja dos modelos de los ocho presentados y trate de describir el contexto de aplicación, alguna finalidad y algún metodológico relacionado con las técnicas e instrumentos a emplear.
- III. El modelo cuantitativo tiene campos de incidencia apropiados, mejor incluso que el cualitativo. Señale cinco ámbitos o dimensiones de una evolución de centros a evaluar según esta concepción.

SANTOS, Guerra Miguel A. “La evolución. Un proceso de diálogo, comprensión y mejora” En: La evolución. Un proceso de diálogo, comprensión y mejora. Granada Ed. Aljiber pp 33-46.

CAPÍTULO II

LA EVALUACIÓN. UN PROCESO DE DIÁLOGO, COMPRENSIÓN Y MEJORA.

LA NATURALEZA DE LA EVALUACIÓN

La evolución educativa es un fenómeno habitualmente circunscrito al aula, referido a los alumnos y militando al control de los conocimientos adquiridos a través de pruebas de diversos tipos. En otro lugar (Santos Guerra, 1988) planteé algunas cuestiones sobre la patología de la evolución educativa, haciendo patentes los problemas de las limitaciones, las desviaciones y las manipulaciones de las que puede ser objeto. Plantea de forma negativa, realizada en malas condiciones y utilizada de forma jerárquica permite saber pocas cosas de cómo se produce el aprendizaje y pocas veces sirve para mejorar la práctica de los profesores, y desde luego, el contexto y el funcionamiento de las escuelas.

Aquí me voy a referir a la evolución de programas educativos, no necesariamente circunscritos al marco de la escuela, la educación tiene innumerables contextos y se desarrolla intencionalmente en muchos ámbitos. La evolución de la que hablo se interroga por el valor educativo de los programas y se plantea como un medio de entenderlos y de mejorarlos.

La evaluación es una parte integrante de los proyectos, no algo añadido al final de los mismos, como un complemento o un adorno que se podrá en funcionamiento si queda tiempo y sí se tienen a bien. Se pregunta por el valor de los programas de las acciones. Es, pues, sustancial al hecho mismo de poner en marcha una experiencia. Porque si se diseña, planifica y pone en funcionamiento será imprescindible conocer qué es lo que se consigue por el hecho mismo de poner la iniciativa en acción, por qué esa y no otra, por qué de ese modo y para esos fines. La evolución producirá diálogo, comprensión y mejora de los programas estén funcionando son buenos, no porque hayan sido implantados con la mejor intención producen los mejores efectos.

Hablaré aquí de la evolución de experiencias o de programas. Pero sólo de un tipo de evolución, no de la evolución en general ni de otros modelos o concepciones. Digo esto porque la confusión semántica que se esconde bajo el concepto aparentemente unívoco de evolución es muy grande. Se habla de evaluación refiriéndose a un proceso de carácter comparativo aplicados a mediciones de

resultados, se llama evolución a fenómenos de rendimiento de cuentas impuestos por la autoridad, se denomina evaluación al análisis diagnóstico realizado mediante instrumentos que permiten cuantificar los datos, se habla de evolución cuando se realiza la comprobación del aprendizaje de los alumnos... plantearé aquí algunas cuestiones sobre un determinado tipo de evolución. No es que sea inaceptable otros enfoques o métodos, sino que ha preferido cerrar el arco conceptual para saber a qué me estoy refiriendo. Hablaré de una evolución que tienen las siguientes características:

- *Independiente y por ello comprometida.* Cuando digo independiente, me refiero a una evolución que no esté sometida, sojuzgada, vendida o simplemente alquilada por el poder, el dinero o la tecnología. y hablo de una evaluación no aséptica, no neutral, sino comprometida con unos principios, con unos valores. El educador debe ser imparcial, pero no significa que sea aséptico. Una postura desinteresada, no comprometida, distanciada del mundo cotidiano y de sus valores es "moralmente deficientes" (house, 1990).
- *Cualitativa y no meramente cuantificable.*, porque los procesos que analiza, cuando se trata de programas educativos, son enormemente complejos y la reducción a números suele simplificar y desvirtuar la parte más sustantiva de los mismos.
- *Práctica y no meramente especulativa.* Quiero decir que la evaluación de la que voy a hablar no tiene por finalidad producir conocimiento académico, realizar informes para publicaciones especializadas y ni siquiera elaborar un conocimiento teórico sobre las experiencias. Tienen por finalidad la mejora de programas a través de su comprensión, a través del conocimiento de su naturaleza, funcionamiento y resultados.
- *Democrática y no autocrática,* la evaluación de la que hablo se pone al servicio de los usuarios, no del poder. Porque la evolución, cuando es utilizada torcidamente produce efectos muy negativos. Los evaluadores se ponen al servicio del poder o del dinero. Y, a través de la evolución, se consigue cortar un programa porque no es rentable políticamente o porque no favorece de forma clara a los más pudientes o porque produce lateralmente molestias o críticas para el poder. En este caso, sería deseable no poner en marcha las evoluciones. Conseguir publicidad y control, alcanzar una mejora de programas exclusivamente destinados a los más favorecidos, poner la evolución al servicio de la justicia, hace de la evolución un instrumento y un proceso de previsión.
- *Procesual, no meramente final.* La evolución que propongo se realiza durante el proceso y no una vez terminado el programa,. Porque es durante

el mismo cuando se puede conocer lo que en él sucede. Y porque durante su desarrollo se puede modificar y mejorar. Realizada al final del mismo, como una apostilla, como un apéndice, pierde la capacidad de generar comprensión de lo que realmente va sucediendo. Y aunque al final se analice todo el proceso, la perspectiva está adulterada por el gradiente de la meta.

- *Participativa, no mecanicista.* La evolución de la que hablo da voz a los participantes, no se realiza a través de pruebas externas y de análisis ajenos a la opinión de los protagonistas. Son ellos los que emiten su valoración sobre el programa, aunque no sea ésta la única voz y la única perspectiva que se tienen en cuenta.
- *Colegiado.* No individualista. Es un tipo de evaluación que asume un equipo y no sólo un individuo-. No porque la que se realiza bajo la responsabilidad única de una persona sea deficiente sino porque la realiza por un equipo goza del aval del contraste, de la pluralidad de los enfoques, de una mayor garantía de rigor, de una diversificación estratégica de acceso y actuación.
- *Externa, aunque de iniciativa interna.* Es decir, son los propios participantes los que la demandan. Y para realizarla requiere la colaboración exterior, por considerar que desde fuera puede tomarse una perspectiva complementaria y puede trabajarse en unas condiciones favorables para conseguir una información veraz.

Cuando no se dispone de evaluadores externos, puede realizarse una Autoevaluación que esté asentada sobre los pilares ideológicos y metodológicos que aquí se plantean.

Hablaré de la evolución no como de una tasación o de una comparación o de un proceso de rendimiento de cuentas sino de la evolución como un proceso de indagación sobre el valor educativo de un programa, de su importancia, exigencias y significados. Para evaluar hay que comprender, dice taxativamente Stenhouse (1984). No hay que olvidar que lo más importante no es el hecho de realizar la evolución, ni siquiera el modo de hacerla sino al servicio de quién se pone. La evolución puede ser utilizada de forma negativa o inutilizada a través de interpretaciones caprichosas e interesadas. Se puede atribuir la valoración de los protagonistas a su falta de buen criterio o de poca exigencia, se puede achacar la mala opinión a deseos malévolos de perjudicar a los responsables... Esta interpretación que preside el uso de la evolución no siempre puede estar controlado por el evaluador, sino que, en parte, es responsabilidad del patrocinador o destinatario de los informes. Existen diversos modos de hacer inútil o perjudicial una evolución. Incluso una evolución que ha sido solicitada por los protagonistas de un programa:

- a.- Esperar que los resultados de la evolución sean elogiosos para los responsables de u programa que demanda voluntariamente una evolución. No es fácil que, sobre todo en los momentos iniciales el informe resulte asimilable en sus vertientes críticas.
- b.- Atribuir las informaciones y las explicaciones poco gratas ala subjetividad del evaluador o a la naturaleza cualitativa de la evolución.
- c.- Demandar de la evolución los juicios de valor que permitan saber a los desrtinatarios y responsables del programa qué es lo que está bien o lo que está mal hecho.
- d.- Exigir a los evaluadores las orientaciones precisas para la mejora de programa, las instrucciones concretas lue provoquen un cambio o las soluciones a los problemas o conflictos existentes.
- e.- Poner los resultados de la evolución al servicio de intereses (políticos, económicos, personales...), encontrando en los evaluadores una excelente excusa para tomar decisiones sin el compromiso de su justificación auténtica.
- f.- Utilizar la evolución como un arma arrojadiza contra quienes piensan o actúan de forma distinta o contraria, principalmente en el, caso de que existan grupos enfrentados en el seno del grupo que desarrolla el programa.

El evaluador puede poner algunos medios para aminorar los peligros del abuso de la evolución. En parte en beneficio de los usuarios de los programas en parte por el interés de la misma evaluación como un proceso potencialmente rico que puede ser desvirtuado por su mal uso. Uno de los modelos de establecer esa garantía es hacer entrega resumida del informe a los usuarios, de tal modo que el patrocinador de la evolución no pueda utilizar impunemente en su provecho. Hacer públicos los informes es un modo de garantizar la democratización de la evolución. El evaluador evitará de esta manera el ser utilizado al servicio de intereses bastardos. De no tomar estas precauciones, el patrocinador podrá usar los servicios profesionales del evaluador para justificar decisiones que de otro modo no se atrevería a tomar o para hacerse publicidad manipulando los contenidos de la evaluación. La servidumbre de los evaluadores hacia el poder que cargo o para la evolución es uno de los riesgos más graves de este tipo de evolución. Si va a ser utilizada en beneficio del dinero, del poder o de la justicia, sería mejor que no se hiciese.

UN PROCESO DE DIÁLOGO, COMPRENSIÓN Y MEJORA.

Algunas de estas ideas han sido expuestas anteriormente en un trabajo conjunto (Angulo y otros, 1991) que es obligado citar aquí. La metáfora del triángulo, parcial y discutible como toda metáfora, nos sirve para reseñar las conexiones que el proceso de evolución, a mi juicio, debe mantener. La evolución supone una plataforma de diálogo entre los evaluadores y los evaluados, entre diversos estamentos de los evaluados, entre los evaluados y las diversas audiencias, entre

éstas y los evaluados, etc. Pero el diálogo tienen una doble finalidad: por una parte, trata de generar comprensión del programa y, por otros de mejorar la calidad del mismo.

La conexión entre las tres dimensiones, que se realizan en diferentes direcciones y sentidos, potencia cada una de las vertientes que aquí aislamos para facilitar el análisis. No son, pues, tres compartimentos estancos y diacrónicamente secuenciados sino partes de un mismo proceso que sólo para el análisis diferenciamos y aislamos. Mientras se realiza el diálogo, y porque se realiza, se produce la comprensión. Cuando se produce la comprensión es más fácil y enriquecer el diálogo. El diálogo fecundo es una parte de la mejora del programa...

La evaluación como diálogo

El juicio de valor que la evaluación realiza se basa y se nutre del diálogo, la discusión y la reflexión compartida de todos los que están implicados directa o indirectamente en la actividad evaluada. El diálogo ha de realizarse en condiciones que garanticen la libertad de opinión que se cimienta en la garantía del anonimato de los informantes y en la seguridad de que la información va a ser tenida en cuenta y utilizada convenientemente.

El diálogo se convierte así en el cambio por el que los distintos participantes en el proceso de evaluación se mueven en busca de la verdad y del valor del programa. Desde la apertura, la flexibilidad, la libertad y la actitud participativa que sustenta un diálogo de calidad se constituye el conocimiento sobre la realidad educativa evaluada.

La evolución así entendida, se basa en la concepción democrática de la acción social. Los destinatarios del programa dan opinión y emiten juicios sobre el valor del mismo.

“Al dar la palabra y adaptar su metodología a la vida diaria de los programas, el evaluador reconoce que su realidad está activamente construida por los sujetos que en ella participan, que en ella participan, y que por lo tanto, ellos mismos pueden cambiar críticamente sus condiciones de vida actuales” (Angulo, 1988).

Ni la verdad ni la valoración correcta están en posesión de personas o grupos privilegiados. Ni de los evaluadores, ni de los implicados en el programa, ni de los usuarios del mismo ni, por supuesto, de los patrocinadores de la evaluación. Es precisamente la conjunción de todas estas perspectivas y opiniones lo que permite profundizar y extender el conocimiento sobre el valor del programa.

Los evaluadores/As no juzgan el programa ni, mucho menos a los responsables del mismo. Facilitan, a través de los datos recogidos, de su interpretación y valoración que sean los propios participantes los que emitan un juicio de valor más

fundamentado y significativo. La evaluación, a través del diálogo que genera entre todos los participantes, se hace preguntas sobre, al menos, los siguientes aspectos:

a.- El sentido y el valor de las realidades que son objeto de la evolución. Objetivamente, no existe una exclusiva visión sobre lo que es valioso desde el punto de vista educativo. De ahí la conveniencia de alimentar el diálogo sobre la cuestión.

b.- El modo de recoger las evidencias que permiten desarrollar la comprensión. Como los datos extraídos de la realidad son la fuente de la valoración es preciso analizar si los manantiales que dan agua a esa fuente están contaminados. De datos recogidos de forma parcial, defectuosa, arbitraria, superficial o tendenciosa, no pueden surgir análisis rigurosos.

La finalidad del trabajo de los evaluadores es orientar la selección de informaciones y puntos de vista que permiten a los interesados ampliar y matizar la comprensión y la reflexión, dialogar con ellos, a través de la negociación inicial de las negociaciones posteriores, de las entrevistas exploratorias y de la discusión de informes y propiciar la interrogación sobre lo que ha de mejorarse a la luz de lo que se ha descubierto. Así pues, el diálogo se alimenta desde niveles estratégicos del proceso:

a.- En la negociación inicial que se pregunta por el sentido de la evolución, por la naturaleza de sus fines, por los focos del análisis, por los métodos de exploración, los procesos de discusión, por las condiciones éticas...

b.- En la exploración que tienen lugar a través de conversaciones informales y de entrevista formalizadas que facilitan y potencian el intercambio y la opinión. Hace preguntas es un modo de avivar el diálogo.

c.- En la negociación de informes que da lugar a la valoración de los puntos de vista sobre el programa y sobre los métodos que han facilitado la realización de esos juicios.

El diálogo se potencia discutiendo los criterios del valor de las acciones educativas, de los resultados que se alcanzan, tanto de los pretendidos como de los secundarios, de la proporción o desproporción de los esfuerzos realizados, del tipo y amplitud del grupo al que va destinado el programa, del análisis de las condiciones en que ese programa se desarrolla...

El diálogo se aviva explicitando los criterios éticos de la evaluación. Con la aplicación de esos criterios se pretende asegurar el ejercicio democrático de la evaluación y de sus resultados, mantener un equilibrio entre el derecho a la privacidad de las acciones y de las informaciones y la necesidad de conocimiento y responsabilidad social, mantener el nivel de justicia democrática en la actividad evaluada.

Los evaluadores han de hacer viable el diálogo desde actitudes abiertas, sencillas, tolerantes y comprensivas. Asimismo, han de buscar las condiciones (tiempo, espacio, explicaciones, garantías...) en que ese diálogo pueda desarrollarse

adecuadamente y han de cumplir unos requisitos éticos que garanticen el respeto a las personas, los intereses sociales y el cumplimiento de las obligaciones asumidas en el marco de la negociación.

Como animadores del diálogo entre destinatarios del programa y entre éstos y los propios evaluadores, éstos han de mostrarse hábiles en la formulación de preguntas, atentos a la captación de las respuestas e inteligentes en la interpretación de la información recibida. Información que no llega solamente a través de los contenidos verbales sin de otras muchas fuentes: la forma de llegar a la entrevista, el papel que desempeña una persona en el desarrollo del programa, el trato que mantienen con el evaluador, el modo de expresarse...

Algunas situaciones de diálogo están formalizadas y se registran de forma inmediata a través de grabaciones magnetofónicas o de registros manuales, pero otras tienen un carácter informal y permiten recabar una información más espontánea aunque más desestructurada. No se debe prescindir de este tipo de comunicación que brinda valoraciones probablemente menos teñidas por el recelo y menos encorsetadas por la disciplina de lo formal.

Son múltiples las plataformas en las que se establece el diálogo y numerosos los niveles en los que tienen lugar y se desarrolla:

- a.- Entre evaluadores y evaluados, que permanecen en un continuo diálogo a través de las entrevistas, de la discusión de los informes, de la interpelación que ambos hacen a la realidad...
- b.- Entre los diferentes evaluados, que pueden contrastar sus opiniones a través del evaluador o directamente entre sí. Es fácil que la evolución da lugar a conversaciones sobre el programa y sobre la forma en que está siendo evaluado.
- c.- Entre la sociedad y las instancias educativas cuando se difunde los informes de La evaluación. El informe, una vez publicado, se convierte en una palestra sobre la que se discute acerca del valor de los programas educativos.

En efecto, la discusión de informes constituye una plataforma para el diálogo. De ahí que será preciso rechazar cualquier tipo de imposición para mantenerlos secretos para hurtar una parte de la información. La publicidad de los informes conduce a un conocimiento más documentado y menos demagógico. Da pie a que se hable del valor educativo de los programas no desde la arbitrariedad y el interés sino desde el rigor de los datos. La sociedad tiene derecho a conocer el valor educativo de las instituciones y de los programas, sobre todo cuando se realizan con fondos públicos. Aceptar y alentar informes privados, secretos o sesgados favorece el engaño, el beneficio privado y la arbitrariedad autoritaria.

Es preciso convertir la evolución y los informes en estructuras sociales de participación y de relación. Un informe no es una conclusión inapelable y taxativa

son un cauce para generar procesos y dinámicas sociales. Los evaluadores, a través de los informes, introducen procesos y demandas democráticas de conocimiento y reflexión.

La evolución como comprensión

La evolución pretende alcanzar un nivel de comprensión sobre el funcionamiento del programa en su contexto, sobre su racionalidad y su sentido educativo, sobre las intenciones educativas que lo han puesto en marcha y sobre los efectos que esté generando.

“Los participantes en el programa actúan de forma razonable según sus circunstancias y oportunidades. La tarea de una evolución es iluminar el raciocinio que dirige el desarrollo de un programa y su evaluación, identificar factores históricos y contextualizados que lo influyen y facilitan el examen crítico de estos aspectos dentro y fuera de la comunidad sobre la cual actúa el programa” (Kemmis, 1986).

La evolución está guiada por el impulso de la comprensión. Se plantea como finalidad entender por qué las cosas han llegado a ser como son. Lo cual hace necesario saber cómo son realmente.

“Las verdades que hay que decir acerca de los programas educativos son verdades sociales,... las pretensiones sobre el programa son defendidas y discutidas en un proceso de debate crítico. Lo que se consideran “verdades” acerca del programa dependerá de la calidad del debate. El debate racional consiste en dar razones y defender razonamientos con información y argumentos.

Una evaluación puede soportar una contribución sustancial a este proceso de crítica. Puede reunir información relevante para el programa, sus objetivos y aspiraciones, y puede someterla a un examen profundo.

Puede obtener, articular y compartir la comprensión de por qué el proceso es como es por referencia a los objetivos de quienes en él participan, los condicionamientos circunstanciales y las oportunidades disponibles, y reflexionar sobre esta comprensión a la luz del contexto más amplio de la experiencia del programa como un todo” (Kemmis, 1986).

La comprensión puede estar a diferentes aspectos del programa.

- 1.- *Pretensiones educativas.* Quién las pone en marcha y por qué. Qué finalidades ocultas o manifiestas de otro tipo están vinculadas a las directamente educativas.
- 2.- *Necesidad.* No sólo considerada en sí misma, sino en contraste con otras posibles necesidades de la sociedad.

- 3.- *Destinatarios.* A quiénes va dirigido y qué modos existen de acceder a él. Tanto en lo que respecta a la información como a los canales por los que se accede a él.
- 4.- *Procesos que pone en marcha.* Qué tipo de actividades contienen y qué actitudes, conocimientos o destrezas desarrolla en los destinatarios.
- 5.- *Resultados durante el mismo programa.* Qué es lo que el desarrollo del programa está consiguiendo en los destinatarios, en los responsables, de forma indirecta, en otras personas.
- 6.- *Resultados a largo plazo.* Aunque durante el desarrollo del programa estos efectos no puedan verse, habrá que pensar en lo que sucederá a largo plazo con ese planteamiento.
- 7.- *Efectos secundarios.* Qué repercusiones de carácter no buscado está produciendo el programa, tanto en los participantes como en otras personas, quizá potenciales usuarias del mismo.
- 8.- *Rentabilidad de los costes.* Hay que preguntarse por la proporción entre el esfuerzo y el costo del programa y su rentabilidad, ya que los recursos son limitados.
- 9.- *Rentabilidad social.* El beneficio del programa, desde el punto de vista democrático, no ha de verse solamente en los actuales destinatarios sino en las posibilidades de beneficio en cascada.
- 10.- *Relación oferta demanda.* Es interesante saber qué tipo de ofertas se realizan y por parte de quién en la organización de programas y quiénes u por qué demandan ese tipo de acciones educativas.
- 11.- *Continuidad en el futuro.* Qué tipo de desarrollo posterior tienen el programa, qué posibilidades de continuidad y de profundización encierra en el futuro.
- 12.- *Contexto sincrónico y diacrónico.* El programa se realiza en un lugar y un momento determinado, con unos medios concretos. Y tienen una historia que permiten comprender cómo son las cosas en él actualmente.

Las respuestas a estas cuestiones dependerá del punto de vista y del cuadro de valores desde el que se parta. (Me refiero ahora al evaluador). La respuesta de los participantes dependerá también de su particular visión y, en ocasiones, de los intereses que están en juego dentro del programa. En el desarrollo de la evaluación es posible encontrar participantes:

- a.- Sin opinión definida respecto a algunas cuestiones o todo el programa, sea porque no disponen de datos suficientes, sea porque se contraponen los elementos de juicio.
- b.- Detractores del programa que buscan datos y desarrollan interpretaciones encaminadas a la formulación de juicios negativos.
- c.- Defensores del valor del programa, unas veces de forma incondicional y poco reflexiva y otras con el apoyo de argumentaciones sólidas y consistentes.

“Los evaluadores deben tener en cuenta las distintas perspectivas acerca de un programa, los juicios de valor que se produzcan, su contexto y su historia. Entonces estarán en posición de conglomerar y refinar las opiniones públicas y privadas por las cuales se hace inteligible el programa y se determina definitivamente su valor. La calidad de la educación puede ser juzgada según la calidad de su contribución al debate crítico, informándolo y mejorándolo” (Kemmins, 1986).

No se puede conseguir la comprensión profunda del valor educativo de un programa convirtiéndolo en simplificaciones que no sólo la empobrecen sino que la distorsionan con la excusa de que simplificando es más fácil establecer conclusiones o emitir juicios. La simplificación hace que la valoración sea más fácil pero también más injusta y más errónea.

No puede alcanzarse la comprensión a través de la simplificación de los números o de las estadísticas. La asignación de números a realidades tan complejas encierra el peligro de la apariencia de cientificidad. Puesto que hay fórmulas. Hay rigor. Puesto que hay mediciones, hay objetividad. Puesto que hay cuantificación, hay claridad. Se pueden utilizar mediciones y procedimientos estadísticos en la evolución, se pueden aplicar cuestionarios que recojan información cuantificable, pero todos sus resultados han de pasar por el tamiz del análisis, por el contraste con otros modos de exploración, por la criba de las interpretaciones.

Para conseguir esa comprensión en profundidad es preciso utilizar instrumentos variados y sensible a la riqueza y complejidad estructural y diversidad de interrelaciones que se producen en la realidad educativa. Es necesario trabajar con instrumentos capaces de captar las valoraciones. Aspiraciones, motivaciones, intereses, interpretaciones... de los protagonistas. No se consigue una información idéntica a través de diferentes métodos. Unas veces porque los interesados deforman las perspectivas movidos por el interés, por el papel que desempeñan en el programa o por cualquier otra causa, otras veces porque la naturaleza misma del método de exploración condiciona la información obtenida. No es lo mismo conseguir información a través de la entrevista a un responsable de un programa que observar directamente el desarrollo de la acción.

En su excelente trabajo, Kemmis (1986) plantea la necesidad de solicitar la opinión de todos los que tienen diferentes tipos de valores en un programa.

“Según el principio de la pluralidad de sistema de valores, los evaluadores deberán conocer y tener en cuenta todos los sistemas de valores en juego, ya que todos son relevantes a la hora de emitir juicios sobre el programa”.

Esto exige la realización de exámenes cruzados de los diferentes métodos. Al contrastar la información recibida a través de dos o más métodos, se depura la

comprensión. Exige e también el contraste de las opiniones de diferentes usuarios, ya que las valoraciones que hacen unos no necesariamente coincide con las que hacen de otros.

La búsqueda de la comprensión tienen como finalidad preguntarse por los procesos y los resultados de los programas, tanto de los que tienen un carácter explícito como de lo que tienen una condición subrepticia, tanto de los directamente pretendidos como de los que aparecen de forma derivada o secundaria, tanto de los relativos a cada uno de los destinatarios como de los de carácter general, tanto de los relativos al conocimiento como a otras esferas del desarrollo personal y social.

La comprensión se realiza a través de un lenguaje accesible, no sólo para los directamente implicados sino para cualquier otro ciudadano interesado en la realidad que se evalúa. No facilitan la comprensión ni el lenguaje tecnicista, ni la abundancia de fórmula compleja ni las conceptualizaciones académicas que, por otra parte, suponen un nivel de codificación que cambian el registro de las informaciones emitidas por los participantes.

La evolución como mejora.

La evaluación no se cierra sobre sí misma sino que pretende una mejora no sólo de los resultados de la racionalidad y de la justicia de las prácticas educativas. No se evalúa para estar entretenidos evaluando, para decir que se está realizando una evaluación, para controlar los programas, para hacer publicidad o para crear conocimiento... Fundamentalmente se hace la evaluación para conseguir la mejora de los programas: del que está en curso y de otros que se pongan en marcha.

“Si la evaluación educativa es una reflexión sistemática sobre la acción docente/disciente y organizativo/funcional tienen que ser la base de las innovaciones que se introduzcan” (Casanova, 1992).

La mejora puede consistir, pues, en diversos fenómenos referidos al programa.

- a.- Por una parte, tienen que ver con la introducción de innovaciones que, desde la comprensión que se a generado, se consideran positivas para la calidad.
- b.- La pregunta misma por la calidad se convierte en una mejora real. Es decir, que la preocupación por conocer qué se entiende por mejora es ya un beneficio que enriquece el programa.
- c.- La participación de los destinatarios en la valoración del programa lleva consigo un factor educativo importante, con tal de que no se conviertan un simple entrenamient6o. Aumenta la participación de los destinatarios del programa en su desarrollo y valoración es, a mi juicio, un modo e mejorarlo.

La mejora habría de ser definida, no tanto a través de unos parámetros establecidos a priori sino en la propia dinámica de la acción contextualizada en la que aparecen un fuerte juego de expectativas, motivaciones, intereses, deseos,

sentimientos y valores. De cualquier forma, habría que superar un planteamiento simplista que sólo atiende a la valoración en la obtención de rendimientos terminales, sobre todo cuando se miden a través de pruebas estandarizadas. No es fácil prescindir del resultado final como uno de los ingredientes del éxito del programa. Pero hay que superar esa visión simplista. Difícilmente podrá considerarse un éxito el programa aquella persona que no supere las pruebas, si las hubiere. Pero no bastaría, a mi juicio, superarlas para decir que el programa ha resultado positivo. Consecuentemente, la mejora, no consistirá solamente en la consecución de mejores resultados finales. Existen otras cuestiones sobre las que habrán de interrogarse quienes piensen en la mejora de un programa:

- a.- La justicia de sus planteamientos desde el punto de vista social y educativo.
- b.- La racionalidad de sus prácticas en cuanto a los fines perseguidos y a los métodos elegidos para su logro.
- c.- La bondad de las relaciones entre los organizadores y los participantes y las de éstos entre sí.

No todo cambio es, por el hecho de serlo, una mejora. Lo que sucede con la evolución es que permite descubrir, a través de la comprensión, en qué consiste la auténtica mejora de los programas.

“La evaluación educativa es un proceso que, en parte, nos ayuda a determinar si lo que hacemos en las escuelas contribuyen a conseguir fines valiosos o si es antitético con estos fines. Que hay diversidad de versiones de valor es indudablemente verdad. Este es uno de los factores que hacen a la educación más compleja que la medicina... Esto hace de la evaluación educativa una tarea difícil y compleja, sin embargo a fin de cuentas algunos valores han de buscarse. Algunos juicios deben ser formulados acerca de lo que se hace o lo que ocurre. No tenemos otro modo de conocer si estamos educando o deseducando” (Eisner, 1985).

La evaluación facilita la mejora al preguntar por el valor educativo del programa, al facilitar la comprensión de lo que sucede en el mismo, al provocar reflexión y el debate de los implicados, al urgir las respuestas sobre los posibles cambios.

La mejora de los programas puede tener un carácter inmediato o bien diferido. En el tiempo se encuentra a veces un obstáculo cuando se buscan efectos rápidos que no pueden darse con inmediatez o cuando se difieren tanto los informes de manea que ya no pueden aprovecharse para la mejora del programa, y que los protagonistas son diferentes o no es posible intervenir adecuadamente.

La alternativa metodológica que supone el estudio de caso en el campo de la valuación vienen da por la riqueza de posibilidades que presentar su capacidad de analizar la realidad educativa en términos de su utilidad,

utilidad inmediata para la acción institucional y para la práctica docente, engarzadas en la variedad de situaciones que generan y en la diversidad de contextos con los que interactúan. Utilidad deferida en cuanto a la progresiva secesión de los mismos y su palatina acumulación (Breznes y Carretero, 1990).

Otra cosa son las estrategias que han de ponerse en marcha para conseguir esta mejora. Los evaluadores no tienen, a mi juicio, la responsabilidad de poner en marcha esas estrategias, pero sí se preocupan porque la evolución no esté abocada a la esterilidad.

La evolución facilita el perfeccionamiento de los profesionales, al hacer que, desde la independencia de las preguntas y de los juicios externos, se interroguen, dialoguen y entiendan lo que sucede con el programa.

La mejora se hace posible no sólo en el programa evaluado sino en otros de carácter similar que ya estén realizándose o que se preparen en el futuro. La transferibilidad permite atender diversos tipos de cuestiones sobre la naturaleza y riqueza educativa en determinadas estrategias de intervención.

La conexión de la evolución con la mejora puede potenciarse si existen unas condiciones determinadas, a saber:

- a.- Si la iniciativa de realizar la evaluación es ampliamente democrática y no está solamente promovida desde el exterior o por sólo una parte de los participantes en el programa.
- b.- Si el proceso de exploración ha sido intenso y ha permitido poner sobre el tapete cuestiones importantes sobre el desarrollo y la valoración del mismo.
- c.- Si están solidamente fundados los juicios que se expresan en la evolución y rigurosamente argumentadas las interpretaciones.
- d.- Si los evaluadores aportan sugerencias sobre la mejora y la viabilidad de las mismas, en el buen entendido de que son los responsables del programa y los participantes en el mismo quienes han de hacer efectivos los cambios.

La evaluación que propongo se siente comprometida no sólo con la mejora sino que toma partido por unos determinados valores. Se interesa por aquellos que nos son usuarios del programa y que no siquiera podrían serlo por su condición social. Por carencia de medios o por ser víctimas de un capricho reparto de los bienes públicos.

LA VIABILIDAD DE LA EDUCACIÓN

Existen dificultades para que estos procesos se pongan en marcha en el marco de nuestras instituciones. No es el menor el hecho de que falta sensibilidad y convencimiento respecto a la conveniencia o bondad de estos mecanismos de comprensión y de mejora. O más aún, de que haya motivos para realizar ese

esfuerzo complementario para reflexionar, entender y cambiar. Es decir, los profesionales tendrían que sentir la necesidad y formular consiguientemente la demanda de poner en marcha estos procesos. En las instituciones públicas, en las que los profesionales gozan de la condición de funcionarios se hacen más difícil el estímulo y el interés. Se añade a este hecho el que algunas personas entienden la evaluación como una amenaza, no como una ayuda. Los evaluadores externos pueden ser considerados por algunas personas no sólo como intruso sino como jueces, que sancionan el comportamiento calificándolo de positivo o negativo.

Poner en acción procesos evaluadores de esta naturaleza resulta económicamente costoso ya que exige tiempos y dedicación intensiva de profesionales especializados. La sensibilidad de los responsables políticos y sociales ante estas cuestiones relativas a la calidad habría de ser mucho mayor para que se dedicase un importante caudal de recursos a preguntarse por el valor educativo de los programas, tanto de los programas, tanto de los institucionales como de los situados en otros ámbitos de acción social.

La formación de evaluadores es otra cuestión determinante en el tipo de evaluación de la que hablamos. Lo cual no significa que, hasta que no existan profesionales dedicados a este tipo de trabajo, no se puedan poner en marcha procesos de evaluación de programas. Los procesos de autoevaluación, aunque complejos porque hace que los propios participantes se interroguen sobre sus ideas, valores y prácticas, tienen un interesante papel en el cambio y la mejora:

“Generalmente la puesta en marcha del proceso de autoevaluación pone en evidencia las distintas perspectivas que se subyacen en las categorías usualmente consideradas unívocas que subyacen en la vida escolar” (Martín Rodríguez, 1988).

Difundir los informes, crear plataformas de discusión y debate sobre los mismos, avivaría la preocupación, el diálogo u la comprensión de las realidades educativas. Poner en marcha muchas iniciativas, seguir desarrollando las ya existentes sin establecer unos procesos de reflexión rigurosa. Sistemática y compartida sobre ellos, encierra el peligro del empobrecimiento, la rutina y la esterilidad .

Generar teoría sobre la evaluación sin atender a las condiciones que la hagan viable, desarrollar u discurso elaborado sobre la importancia de la evaluación sin preocuparse de las estrategias que la hagan deseable y posible, establecer desde instancias públicas prescripciones sobre la evaluación sin atender las exigencias de su implantación y difusión, crea un doble, y a veces contradictorio, discursos que desorienta y que irritan a los profesionales que está comprometidos en los programas.

CASANOVA, María Antonieta; "Evaluación: concepto, tipología y objetivos" En: La evaluación educativa. España. Biblioteca Normalista. Pp 67-102

CAPÍTULO III

EVALUACIÓN: CONCEPTO, TIPOLOGÍA Y OBJETIVOS

EL CONCEPTO COMO BASE DE LA ESTRATEGIA DE ACTUACIÓN

Como ya quedó indicado en el capítulo I hablar, en general, sobre la evolución del concepto de evaluación, éste se ha ido transformando y ha ido incorporando nuevos elementos a su definición según se ha profundizado en su sentido, sus aplicaciones y las virtualidades que ofrece en los diferentes ámbitos de aplicación, que también han quedado señalados.

Ciñéndonos a los procesos de enseñanza y aprendizaje y, más en concreto, a la evolución de los aprendizajes que alcanza el alumnado, la concepción de la misma también ha variado con el paso del tiempo y el avance de las teorías de la psicología evolutiva y la psicología del aprendizaje.

Desde sus comienzos, la evolución aparece influida por su procedencia del campo empresarial. Por eso, al igual que los empresarios miden cuantitativamente los resultados de su producción, en el campo educativo se pretendió medir el proceso del alumno cuantificando lo aprendido. Ello hace que se equipare a "medida" y que durante muchos años (demasiados, ya que en ocasiones llega hasta nuestros días) lo que se intente al evaluar es medir la cantidad de conocimientos dominados por los alumnos. Las cabezas bien llenas a las que aludían Montaigne han resultado objeto principal de la evolución de aprendizajes. La aparición de los tests y las escalas graduadas a principio del presente siglo¹, absolutamente cuantificadas y automáticamente aplicables, contribuyeron a dar carta de naturaleza a esta concepción evaluadora. Pedagogos prestigiados y con una absoluta profesionalidad y categoría educativa, adoptan durante largo tiempo este concepto, aun intentando claramente que su incidencia no resultara negativa para la formación de los niños o jóvenes, ya que el examen como única prueba y el número como expresión del

¹ Rice es el primer psicólogo que aplica una prueba formal para medir la ortografía de 30.000 escolares (1897-1898) y Thorndike elabora la primera escala graduada para "la escritura a mano de los niños" en 1900, que se publica en 1912. siguieron los tests y escalas de Courtis, Ayres, Hillegas, Buckingham y otros, sobre materias tales como aritmética, escritura, composición y ortografía. Paralelamente, en 1905 aparece la escala de inteligencia Bidet-Simón, revisada en 1908 y 1911, con el concepto de edad mental (versión de 1908). Su adaptación a Estados Unidos data de 1916: Escala de inteligencia Stanford-Binet, cuya versión de Terman- que incorpora el concepto de cociente intelectual es la más aceptada y difundida.

resultado se prestan, sin duda, a numerosas arbitrariedades ya falta de rigor en su aplicación; amparándose además, en la aparente objetividad del número para encubrir su impropio modo de valorar. La cita que reproduzco a continuación, de Alejandro Galí (1934, 34 y 37). Es expresiva de esta situación, en exceso generalizada: "Otros sistemas de control también equívocos, son las puntuaciones y los exámenes. En el primero de dichos sistemas se han hecho verdaderas de combinación y cálculo que tienen toda la apariencia de una determinación objetiva de conocimientos. *Nada más lejos de la realidad.* Hay que tener presente que la base de cálculo de las puntuaciones son las notas dadas por los maestros de una manera fundamentalmente arbitraria, no como pudiera suponerse por mala fe, sino por falta de elementos de referencia fidedignos ¿Por qué una lección o un ejercicio vale siete puntos, y no ocho o seis? ¿Es que el maestro puede dar fe de un exacto y positivo trabajo de contraste improvisado en un instante? (...) en este siete o en este ocho que se quieren adjudicar con tanta sensatez y justicia influyen, a despecho de uno mismo, mil circunstancias de temperamento personal, variable a cada momento que pasa, y 7 mil aspiraciones y tendencias latentes que atraen y repelen los puntos del mismo modo que el polo magnético la aguja imantada. (...) Los exámenes, aún, más que las puntuaciones, pueden dar una errónea sensación de objetividad. De ellos se puede decir como de aquel famoso individuo que de nos el borracho, jugador, pendenciero y holgazán, hubiera sido una excelente persona". No obstante, téngase en cuenta que, a pesar de estas reflexiones, la obra de Galí se titula *La medida objetiva del trabajo escolar*. En la misma obra se observa claramente esa influencia empresarial trasvasada a la educación en algunas de sus expresiones: "El maestro puede decirse que no ha dado un paso en el deseo de precisar las ganancias de un negocio, tan apreciable al menos como el del comerciante" (1934, 29).

En fin, que la "evolución" (término y concepto de aparición reciente en el campo de la pedagogía, pues los trabajos de este siglo, como antes ha quedado apuntado) ha sido interpretada como sinónimo de "medida" durante el más largo periodo de la historia pedagógica, y es en los tiempos actuales cuando está variando su concepción, en función de los avances que experimentan las ciencias humanas, las concepciones diferentes que se poseen en relación con la formación de la persona y con los modos peculiares que ésta tiene de aprender, evidentemente, por último y desde una óptica estrictamente educativa, con la intencionalidad clara de aprovechar al máximo la evaluación para optimizar los procesos educativos.

En la actualidad se hace patente una divergencia entre los conceptos evaluación que se manejan a nivel teórico y la práctica real en las aulas. Creo que una buena parte de los profesionales que nos dedicamos a la educación estamos de acuerdo en la necesidad de incorporar a los procesos de enseñanza un modelo de evaluación cualitativo, que sea capaz de ofrecer datos enriquecedores acerca del desarrollo del alumnado y no sólo de los resultados que obtienen a través de

medios no precisamente muy fiables, como más adelante comentaremos y que ya anticipaba Alejandro Galí. Como consecuencia, una evolución que constituya un elemento curricular más que ayude a mejorar todo tipo de aprendizajes. Esta postura se hace palpable en R.S y Rein, M.: 1972; Parlett, M. y Hamilton, D.: 1976; Guba, E.G.: 1978; Fernández Pérez, M.: 1988; Rosales, G.: 1990; Santos Guerra, M. A.: 1993) y en el intercambio de opiniones y experiencias con numerosos profesores. No obstante, parecen existir dificultades importantes para llevar a la práctica ese modelo de evolución, bien sean éstas por la exigencia de cambio de mentalidad que implican, por la presión que el modelo social ejerce sobre el educativo o por la necesidad de otras modificaciones estructurales y organizativas del sistema escolar que no se producen y que favorecerían su adecuada aplicación.

El problema de su incorporación al quehacer en el aula proviene, precisamente, de que no supone sólo adoptar un nuevo concepto de evaluación, estar de acuerdo con él en un plano meramente intelectual, sino que implica cambiar las prácticas que se llevan a cabo en las aulas e invertir, en muchos casos, sus valores. Los alumnos estudian para aprobar. Los profesores enseñan para que sus alumnos superen las evaluaciones. Lo que tienen valor real en la enseñanza es lo que se evalúa, de lo contrario, nadie lo tiene en cuenta. Las familias se preocupan de la situación de aprendizaje de sus hijos cuando éstos reprueban, en fin..., serían innumerables los ejemplos que avalan la importancia –erróneamente entendida- de la evolución en los procesos de enseñanza y aprendizaje y cuyo sentido resultaría imprescindible modificar para que esa importancia le viniera dada por su contribución a la mejora de esos mismos procesos y de sus resultados, como consecuencia de una aplicación como elemento de poder o de mantenimiento de la disciplina, no como instrumento para la promoción u obtención de un título, no como exclusivo factor de comprobación de lo que se “aprende”, nunca como fin de la educación (que es lo que resulta ser en muchos casos para demasiados alumnos, profesores, padres o directivos). No se enseña para “aprobar”. Se enseña y se aprende para alcanzar una plena e integral formación como persona. Ya ese objetivo fundamental debe contribuir toda la organización del sistema educativo institucional. También debe hacerlo la evolución, como elemento central de los procesos reales de enseñanza que ocurren cada día en la comunicación entre jóvenes y adultos y que, como probaremos antes en una cometa enumeración de ejemplos comunes, llega a desfigurar las metas y el camino que en principio conforman el proceso educativo.

Las definiciones que ahora se manejan de evaluación son muy similares, pues, como antes decía, teóricamente existe una gran coincidencia acerca de su papel en la educación. No obstante, voy a formular el concepto de evolución del que parto, porque es la base en la que nos apoyaremos para desarrollar, a continuación, su estrategia de aplicación, si bien especificando más adelante la

diferente topología valoradora con la que es posible contar y que, según las situaciones, será más o menos útil.

La evolución aplicada a la enseñanza y el aprendizaje consiste en un proceso sistemático y riguroso de obtención de datos, incorporando al proceso en un proceso educativo desde su comienzo, de manera que sea posible disponer de información continua y significativa para conocer la situación, formar juicios de valor con respecto a ella, y tomar las decisiones adecuadas para proseguir la actividad educativa mejorándola progresivamente.

Como es evidente, en la misma definición ya estamos orientando la evolución y su modo de desarrollo hacia los procesos y su funcionalidad formativa.

Si conceptualizáramos la evolución de un modo más general y para cualquiera de las funciones que puede desempeñar en el campo educativo, se podría definir como *una obtención de formación rigurosa y sistemática para contar con datos válidos y fiables acerca de una situación con objeto de formar y emitir un juicio de valor con respecto a ella. Estas valoraciones permitirán tomar las decisiones consecuentes en orden a corregir o mejorar la situación evaluada.*

En este segundo planteamos, la evaluación aplicada puede ser puntual y sumativa, con el único objetivo de contrastar el conjunto de informaciones alcanzadas con los criterios referenciales de evolución, es decir, con la finalidad de comprobar lo conseguido o su mayor o menor mérito ante lo que pretendía en el proyecto inicia.

Así, vemos que el concepto de evaluación del que se parta condiciona el modelo de desarrollo de la misma, pues si su meta y su funcionalidad son eminentemente formativas, todos los pasos que se den para ponerla en práctica deben serlo igualmente. Es decir, el planteamiento debe iniciar en los fines que se pretenden e incorporar al proceso a todos los implicados en la realización o afectados por los resultados que pueden aparecer; las técnicas e instrumentos que se determinen y seleccionen o elaboren deben contribuir también a la mejora del proceso evaluado, el informe final y las decisiones que se tomen colaborarán a su positiva aceptación por parte de los destinatarios y éstos, del mismo modo, serán todos los interesados e integrantes de las actividades o situaciones evaluadas. Por el contrario, si la finalidad de la evolución es sumativa, tanto el planteamiento inicial como las técnicas e instrumentos utilizados deberán ser válidos y útiles para permitir valorar los productos o resultados que se evalúan.

Como el objetivo de esta obra es realizar una propuesta para la evaluación de los procesos de enseñanza y aprendizaje, la definición que mejor se ajusta a nuestro propósito es la primera, que no se opone a la segunda, pero sí matiza y se atiene a la peculiaridad de evolución que vamos a desarrollar. Por tanto, sirva esta observación para dejar claro que, de ahora en adelante –siempre que no se

especifique nada en contra-, el modelo de evolución al que nos referiremos es fundamentalmente cualitativo y formativo, que es el que consideramos válido y adecuado para evaluar los procesos de formación humana y de cualquier situación en la que lo humano sea prioritario. Es el modelo que ofrece mayor riqueza de datos útiles para comprender en toda su amplitud y profundidad el proceder de las personas y que permite, por la posibilidad de intervenir y perfeccionar sus desenvolvimiento o actuación. En este sentido comenta Stufflebeam: "El propósito más importante de la evolución no es demostrar, sino perfeccionar..." (Stufflebeam. D.L. y Shinkfield, A. J: 1987,175), afirmación que comparto por entero.

Esta aplicación de la evolución para la mejora de procesos debe contribuir, por parte, al cambio de imagen que, en general, se tienen de la evolución y de los fines que persigue. Me explico: con los conceptos de evolución manejados habitualmente, que han incidido sobre todo en la comprobación de resultados para su calificación posterior -positiva/negativa, aprobando/reprobando, admitido/rechazado...-. la interpretación que se viene haciendo de sus funciones se limitan a entenderla como instrumento sancionador o como arma de poder para quien evalúa. Esto conduce, lógicamente, a que cuando se propone llevar a cabo la evaluación de los centros educativos o del profesorado –por citar un ejemplo- éstos presentan serios reparos a la misma y muestran gran reticencia para presentarse a la actuación evaluadora. ¿Por qué? Porque no se hace patente su utilidad ni sus aspectos positivos favorecedores de una mejora palpable para la profesionalidad de los evaluados, sino que solamente se percibe su carga controladora, sancionadora, etc.

Por otro lado, también es cierto que cuando evaluamos solemos hacerlo para detectar lo negativo, -errores de aprendizaje, fallos de funcionamiento, etc.- y, en general, destacamos escasamente lo positivo. Creo que ambos aspectos tienen importancia como para ser evaluados y puestos manifiestos pero, en el momento de informar acerca de los resultados de una evaluación o de la evolución de un proceso continuo de valoración, habría que comenzar dando especial relieve a todo lo que se ha evaluado muy positivamente o, más en concreto, a los aprendizajes que los alumnos han alcanzado de modo definitivo, como seguridad en su utilización, etc., procedimiento que, por una parte, posee un fuerte ingrediente motivador para que el estudiante siga trabajando y que, por otra, ofrece pautas al profesor de lo que resulta interesante para sus alumnos y, por tanto, del camino por el que puede/debe continuar su programación. No obstante, también es imprescindible detectar los errores de aprendizaje de cada alumno para poder subsanarlo de modo inmediato, sin esperar a efectuar un "control", "examen"...(llámese como se quiera), al cabo de un mes, cuando pocas o ninguna solución eficaz tiene ya el remedio que quiera aplicarse.

Pero si estamos empeñados en que cambie (para mejor) la imagen y el sentido de la ablución, la clave está en:

- a) Detectar el error de aprendizaje en el momento en que se produce, de manera que surta efectos para la aclaración de determinadas cuestiones ni comprendidas adecuadamente y el alumno pueda continuar avanzando en su formación sin rémoras por conceptos mal adquiridos, procedimientos no utilizados o actitudes negativas en el grupo o frente al trabajo.
- b) En consecuencia, ese error detectado no tienen efectos sancionadores, puesto que de él no se deriva una "calificación" negativa, cosa que sí ocurriría así se comprobara en un examen. El error, en ese último supuesto, no sería ya una llamada de atención para superar una disfunción de aprendizaje, sino que se convirtiera en un elemento para emitir un juicio negativo de ese alumno en relación con los objetivos pretendidos. .

Si hacemos a la autoevaluación de la práctica docente –por ampliar el argumento anterior-, hay que insistir en la necesidad de evaluar lo positivo por delante de lo negativo. Un ejemplo: si durante el desarrollo de una unidad se detecta que ciertas actividades no son bien aceptadas por el alumno, el único dato que nos aporta esta valoración es que debemos cambiarlas, pero nunca nos dice en la unidad y comprobamos procesualmente qué tipos de actividades resultan más interesantes, mas eficaces para alcanzar determinados objetivos, etc., sí obtenemos datos claros y precisos para saber cuál es la vía más eficaz para nuestro trabajo posterior, por dónde debemos encauzar la programación para llegar a los resultados previstos.

En definitiva, en cambio de imagen de la evaluación debe ser la consecuencia de su cambio real de contenido y funcionalidad con la que se aplica. Si se logra modificar el concepto y los efectos que la evolución ha venido teniendo hasta tiempos tan recientes como hasta ahora mismo –en demasiados casos-, está debería no sólo generalizarse en los centros en un plazo de tiempo razonable, sino que debería incluso ser exigida. Y por lo que afecta a los procesos de aprendizaje, tendría que ser aceptada favorablemente por el alumno, al convertirse en un elemento curricular más – y muy importante – que, de forma evidente, contribuyera a su mejor formación y desarrollo, y en la que, además, muchos a alumnos intervinieran –fueran tenidos en cuenta- para decidir la valoración final de lo que han alcanzado.

En cualquiera de los casos que hemos comentado y en las dos definiciones que proponemos –según su aplicación más o menos general-, queda claro que todo proceso evaluador debe seguir unas fases que lo caracterizan y sin, las cuales, no se puede hablar de evaluación en sentido estricta. Estas frases se concretan en:

- a) Recopilación de datos con rigor y sistematicidad.
- b) Análisis de la información obtenida.

- c) Formulación de conclusiones.
- d) Establecimiento de un juicio de valor del objetivo evaluado.
- e) Adopción de medidas para continuar la actuación correctamente.

Estos pasos tendrán una duración mayor o menor en función del objetivo evaluado y de las metas que se hayan propuesto para la evaluación: si se evalúa el funcionamiento de un departamento didáctico, será preciso al menos, un plazo de tres meses para observar su proceso de desarrollo y los resultados que se van consiguiendo. Pero si se trata de evaluar el proceso de aprendizaje de un alumno o alumna, estos pasos pueden darse guante un periodo de clase. Un ejemplo: el profesor observa la realización de determinadas actividades por parte de un alumno, detecta si la s realiza correcta o incorrectamente y, de acuerdo con el análisis hecho y el juicio de valor que se ha formado, refuerza o amplía esas actividades o da alguna explicación complementaria y cambia las actividades por otras más sencillas. En ambas situaciones puede darse estrictamente el proceso evaluador enumerado en fases, pero en cada una presenta características diferentes, pues muy diferentes son las realidades de los hechos educativos.

Aunque parece razonable el anterior planteamiento, que siempre supone llega a una valoración del objeto evaluado, entre las tendencias evaluadoras existentes a nivel internacional encontramos a los denominados "descriptores", que establecen el papel de la evaluación en una descripción o enumeración de los hechos a tal y como se producen; abogan porque ésta resulte lo más amplia, profunda y enriquecedora posible, pero limitan su función a exponer lo observado, simplemente, los datos obtenidos por cualquiera de las técnicas seleccionadas para llegar a ellos.

Es una postura que, desde mi punto de vista, identifica evaluación con investigación,, y en la cual se pierde la peculiaridad de cada una de estos dos actuaciones. En la figura 5 presento un paralelismo entre ambos procesos, donde se comprueban los ámbitos que resultan comunes y los que son diferenciadores. Lo característico de la investigación es partir de una hipótesis de trabajo que, mediante el estudio posterior, se ratifica o se desestima, por lo tanto, su función finaliza (al menos, teóricamente) con un informe descriptivo en el que se da cuenta del proceso realizado y de las conclusiones obtenidas. La evolución, por su parte, no arranca de una hipótesis previa, sino de la decisión de evaluar unos determinados objetos, procesos, situaciones, etc., sigue un camino similar y paralelo al de la investigación y no finaliza hasta valorar lo evaluado y proponer o tomar las medidas necesarias para corregir o reforzar la línea de actuación. Por lo tanto, coinciden en las fases intermedias y se diferencian en el comienzo y final del trabajo.

Figura 5
Investigación y Evaluación

Fases de la investigación	Fases de la valuación
Definición del problema	La misma
Planteamiento de la hipótesis	xxxxxxxxxxxxx
Planificación: diseño, técnicas, instrumentos, tiempos, agentes, destinatarios.	La misma
Ejecución: recopilación, análisis y tratamiento de los dato	La misma
Informe: descripción y conclusiones	La misma
xxxxxxxxxxxxxxxxxxxxx	Valoración del objeto evaluado
xxxxxxxxxxxxxxxxxxxxx	Toma de decisiones

Éste es el motivo de que, según los objetivos que se pretendan, sea necesario plantear una investigación o una evaluación para cada situación o problemática específica que se presente en la educación. La propiedad de aplicar una u otra estará en función de la utilidad que preste para su resolución más adecuada.

Popham, W. J. (1980, 20-22) establece las diferencias entre investigación y evaluación educativa, fundamentalmente, en tres cuestiones:

- *Definición del problema:* "Los investigadores quieren sacar conclusiones. A los evaluadores les interesan más las decisiones (...) Quizá en el mundo real no se encuentren nunca estas diferencias de un modo tan marcado como el que aquí se presenta. Difícilmente puede haber un investigador básico que desdeñe una información que se pueda utilizar para tomar decisiones. Y, asimismo, será raro encontrar un evaluador que no esté interesado en entender un fenómeno por el mismo hecho de entenderlo. Lo que sucede es que los investigadores tienden a enfocar sus trabajos con vistas a derivar conclusiones, los evaluadores tratan de plantearse los temas con vistas a posibilitar mejores decisiones".
- *Posibilidad de generalizar los resultados:* "Un resultado ideal para una investigación, serían aquellos descubrimientos que pudiesen ser generalizados a una amplia variedad de situaciones comparables. (...) La evaluación, por el contrario, se centra en un fenómeno educativo en particular. No se pretende generalizar los resultados a otras situaciones. El interés se centra en esta situación y en las decisiones a tomar sobre ella".
- *Papel de la valoración en la investigación:* "El evaluador está obligado a valorar un fenómeno educativo para poder tomar decisiones. Los investigadores, por su parte, buscan la verdad científica sin ningún deseo de incorporar estimaciones de valor a sus hallazgos".

La evolución del concepto de evolución que comentamos en el primer capítulo se corresponde, efectivamente, con esta funcionalidad que ofrece la evaluación para tomar decisiones y, por ello, para mejorar los procesos y/o los resultados de las situaciones valoradas.

Esta posición de intento identificador entre investigación y evaluación se da cuando se aborda el estudio de los sistemas educativos o de las instituciones escolares. Resulta prácticamente invariable cuando aplicamos el término/concepto evaluación a los procesos de enseñanza-aprendizaje, ya que si no se llega a determinar el modo y el grado de aprendizaje del m alumnado no es posible hacerle progresar en su formación. El profesor investiga, evalúa y decide el camino que debe seguir en cada momento de modo continuo, en interacción permanente con sus alumnos. El alumno se forma., igualmente, mediante la autoevaluación de su propio trabajo, que le ofrece datos para elegir la vía más idónea por la cual continuar en su estudio para alcanzar la meta prevista (Lofficier, A,: 1994).

TOPOLOGÍA DE LA EVALUACIÓN

Por clarificar la terminología y conceptos que ésta llena implícitos, y que van aparecer constantemente a lo largo de las páginas siguientes, en la figura 6 presentamos una topología de la evolución que creemos útil metodológicamente para situarnos en este campo y para mostrar las diferentes posibilidades con las cuales podemos aplicar la evaluación y las virtualidades que se obtendrían con ellas.

Comentamos a continuación, con cierto detenimiento, cada uno de los tipos que aparecen en la figura citada.

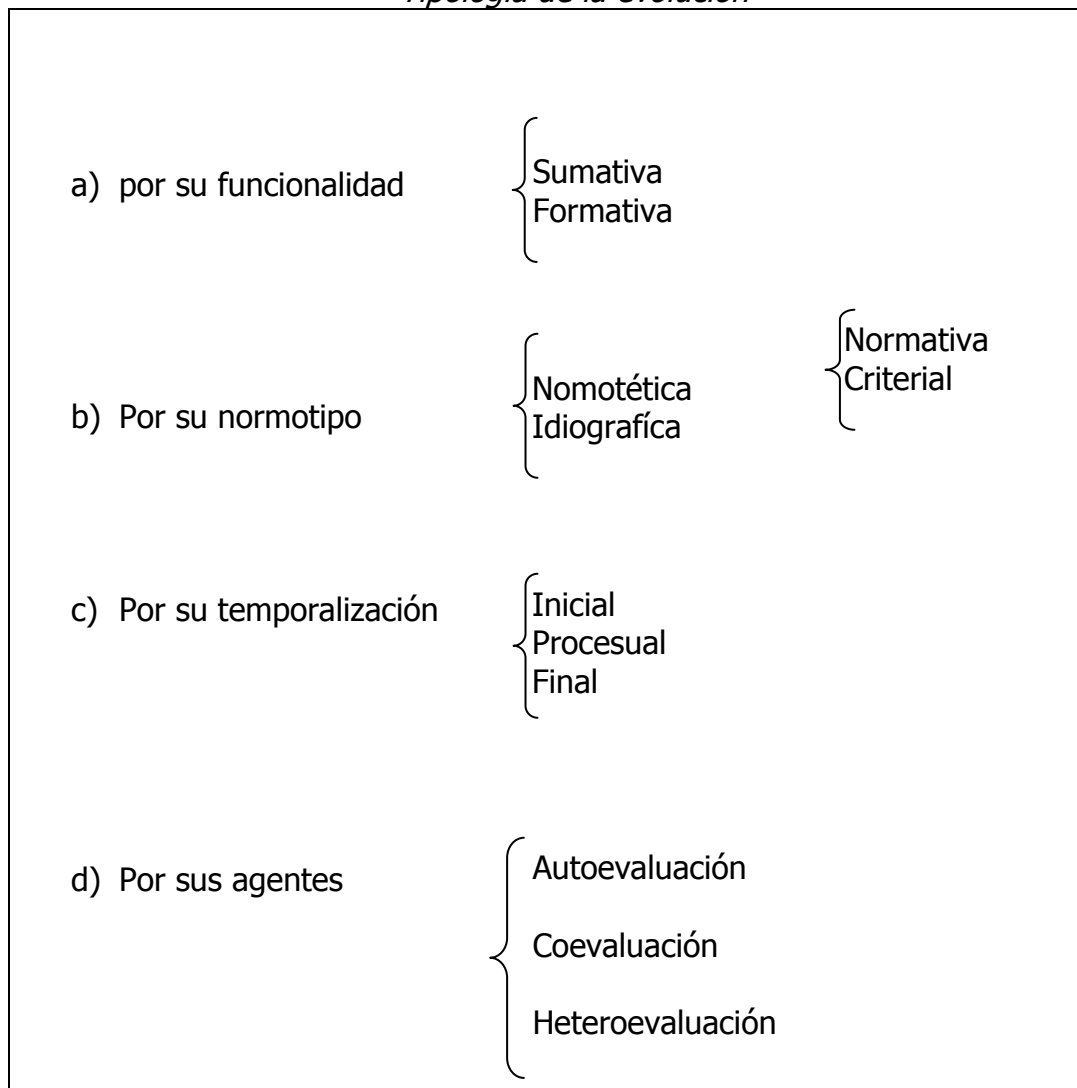
La evolución según su funcionalidad

Las funciones que se asignan o se pueden asignar a la evaluación son diversas. Algunos autores distinguen un buen número de finalidades que es posible alcanzar mediante su aplicación, y de acuerdo con ellas determinan para la evaluación funciones tales como la predictiva, de regulación, formativa, prospectiva, de control de calidad, descriptiva, de verificación, de desarrollo, etc. (De Ketele, J. M. Y Roegiers, X.: 1994, 121-126). Efectivamente, es altivamente fácil hacer diferenciaciones sutiles entre unas funciones y otras, según la finalidad principal que se pretendida conseguir, y realizar clasificaciones amplias. No obstante, aquí nos vamos a centrar solamente en las dos funciones principales de la evaluación. La sumativa y la formativa, descriptas en su momento por Scriben. M. S.L (1967), por considerar que las otras distinciones que se hacen resultan, en definitiva, aplicaciones concretas de unas de estas dos.

No obstante, cabe aclarar, incluso, que en las aplicaciones prácticas de las dos funciones que abordamos a la evolución de los procesos de enseñanza y aprendizaje, se dan interferencias entre ambas, pues es realmente difícil separar

las actuaciones en la vida de modo estricto, para saber y distinguir cuándo estamos persiguiendo una finalidad u otra. Casi siempre se intenta y o alcanzan varias finalidades conjuntamente, con lucha se están cumpliendo varias funciones de forma simultánea. Nos hallamos inmersos, como dice House, R. (1992, 46), en un "espíritu de encumenismo metodológico", y algo parecido está ocurriendo cuando intentamos encuadrar la evolución en una sola tipología, especialmente en su fase de aplicación práctica. Conviene, a pesar de ello, hacer la distinción clara entre las dos funciones, porque hay situaciones en las que se aplican en su sentido más estricto y con todas las consecuencias que en todas las fases tienen cada una de ellas, por un lado, y porque la selección de una u otra para aplicar en la enseñanza es determinante de todo el proceso interno de funcionamiento que se organice en las aulas.

FIGURA 6
Tipología de la evolución



Función sumativa de la evolución

La funcionalidad sumativa de la evolución resulta apropiada para la valoración de productos o procesos que se consideran terminados, con realizaciones o consecuencias concretas y valores. Su finalidad es determinar el valor de ese producto final (sea un objeto o un grado de aprendizaje), e decidir si el resultado es positivo o negativo, si es válido para lo que se ha hecho o resulta inútil y hay que desecharlo. No se pretende mejorar nada con esta evolución de forma inmediata –en sentido estricto, ya no es posible–, sino valorar definitivamente². Se aplica en un momento concreto, final, cuando es preciso tomar una decisión en algún sentido.

Como ejemplos de las aplicaciones que se hacen habitualmente de esa función evaluadora en los centros docentes tenemos:

- a) la selección de libros de texto. El profesorado los examina, comprueba hasta qué se ajustan a su programación, valora si le resultará útil para su trabajo y el de sus alumnos..., y toma la decisión de comprobarlo o desestimarlos. Se realiza la evaluación de un producto ya acabado, sin posibilidades de modificación inmediata, se lleva a cabo en un plazo corto de tiempo, en un momento concreto, se valora positiva o negativamente y se toma decisión oportuna.
- b) La titulación del alumnado al finalizar una etapa educativa: Después de un número de años determinando en los que se han marcado unos objetivos de aprendizaje para los alumnos y alumnas de la etapa educativa que conduce a un título escolar –y pueden/ deben haber transcurrido con una evaluación formativa y permanente de sus aprendizajes.– se valora si se han alcanzado convenientemente por parte de cada estudiante y, en función de los logros obtenidos, se decide si posee el nivel suficiente para el título o no. En estos momentos ya no es posible seguir realizando actuaciones de formación para la continuación adecuada del aprendizaje; hay que tomar una decisión sobre lo conseguido y el grado de formación alcanzado hasta ahí. La evaluación sumativa aplicada a la valoración de logros por parte de las personas tiene exponentes claros, también, en los exámenes de una “oposición”, o en la prueba de selectividad para acceder a la Universidad, por citar otro ejemplo. Son numerosas las ocasiones en que se aplica, aunque no sea de forma tan absolutamente estricta.

Como puede comprobarse, no es la evolución adecuada para aplicar al desarrollo de procesos, sino que es la apropiada para la valoración de resultados finales. No obstante, la fórmula del examen como medio prácticamente único de evolución, se parece bastante a la aplicación procesual de varias valoraciones sumativas a lo largo de un curso (evaluación continua = exámenes continuos). Lo cual es un error educativo grave que nos ha llevado y nos mantiene en la concepción de la

² Con estas evaluaciones es posible tomar medidas de mejora, pero a largo plazo.

evolución como instrumento comprobador, sancionador y de poder. Dado que el aprendizaje nunca es una simple acumulación de datos, hábitos, destrezas, etc., sino la construcción autoalimentadora de una inteligencia crítica y creadora que supone, en todos los casos, y una progresiva modificación de conductas, la evolución sumativa -frete a la continua y formativa- sólo registraría la fragmentación artificial y nefasta de un continuo, es como querer evaluar la calidad de las aguas de una corriente deteniendo el río d cada cien metros. Los aprendizajes no se suman a otros. Se reorganizan unos con otros, se apoyan, reestructura el saber y el hacer del sujeto, conforman el ser que es cuando ha aprendido a serlo. Un proceso no se debe evaluar como si estuviera compuesto de aparados estancos. Es un continuo y hay que adecuar la evaluación a sus características, nunca a la inversa, por el simple hecho y se acepta socialmente sin reflexión alguna. Se puede entender que personas ajenas a la educación asuman el examen puntual como único medio de comprobar el aprendizaje, pero no es admisible en un profesional de la educación que, como tal, tienen la obligación de dominar su profesión y poner al servicio de ella todos sus conocimientos.

Por estas razones, la evaluación sumativa no va a centrar especialmente nuestra atención a lo largo de la obra, aunque se deberá tener en cuenta para aplicar en los momentos señalados como finales de etapa o en otras situaciones de evolución en que resulte apropiada por sus virtualidades.

Función formativa de la evolución

La evolución con funcionalidad formativa se utiliza en la valoración de procesos (de funcionamiento general, de enseñanza, de aprendizaje...) y supone, por lo tanto, la obtención rigurosa de datos a lo largo e de ese mismo proceso, de modo que en todo momento se posea las decisiones necesarias de forma inmediata. Su finalidad, consecuentemente y como indica su propia denominación, es mejorar o perfeccionar el proceso que se evalúa. Este planteamiento implica que hay que realizar la evolución a lo largo del proceso, de forma paralela y simultanea ala actividad que se lleva a cabo y que se está valorando -nunca situadas exclusivamente al final, como mera comprobación de resultados-. Así, en concreto, teniendo datos y valoraciones permanentes acerca de los aprendizajes que va realizando el alumno y u modo particular de hacerlo (ritmo, estilo, inconvenientes, etc.), en el momento en que surge una disfunción o especial dificultades es posible poner los, medios didácticos adecuados para que pueda superarla sin inconvenientes mayores. Del mismo modo, es fácil detectsar los tipos de actividades o situaciones educativas que favorecen su aprendizaje, para potenciarlas con ese alumno y favorecer, así, su formación más idónea. Como sus efectos son permanentes, puesto que las decisiones se toman también de forma continua, permite una acción reguladora entre el proceso de enseñanza y el proceso de aprendizaje, de manera que son sea exclusivamente el alumno el que deba adaptarse al sistema educativo que se le impone, si no que también ese sistema

educativo se adecuó a las personas que atiende y por las cuales tienen sentido, y ésta desarrollen así sus capacidades totales al máximo. Dicho de otro modo. Mientras que la evaluación sumativa debe asegurarse que el producto evaluado responde a las características del sistema, la evaluación formativa debe garantizar que los medios de ese sistema son adecuados a las características de los implicados en el proceso evaluado (figura 7). La función reguladora que toda evaluación posee se pone aquí de manifiesto en sentido inverso: en el primer caso, es la persona la que debe ajustarse al sistema, y en el segundo es el sistema el que debe adaptarse a la persona. Esta meta última será muy difícil de alcanzar si se continúa utilizando la evolución únicamente como comprobación o como obstáculo que hay que superar, y no se aplica para contribuir a la mejora de los propios procesos educativos: la enseñanza del profesor y el aprendizaje del alumno.

Figura 7

Evolución sumativa y formativa

Evolución formativa	Evaluación sumativa
Es aplicable a la evolución de procesos.	Es aplicable a la evaluación de productos terminados
Se debe incorporar al mismo proceso de funcionamiento como un elemento integrante del mismo.	Se sitúa puntualmente al final de un proceso, cuando éste se considera acabado.
Su finalidad es la mejora del proceso evaluado	Su finalidad es determinar el grado en que se han alcanzado los objetivos previstos y valorar positiva o negativamente el producto evaluado.
Permite tomar medidas de carácter inmediato	Permite tomar medidas a medio y largo plazo.

Este *ajuste* de procesos puede darse por un rasgo importante que interviene en la misma definición de evaluación formativa: su carácter de continuidad. Es decir, la evaluación formativa debe ser continua para poder llevar a cabo su función. Un esquema de su posible aplicación aparece en la figura 8. Como se ve, una reflexión constante sobre el quehacer docente, planteándose siempre cómo se está produciendo el proceso de enseñanza y aprendizaje, si va funcionando como estaba previsto, si es necesario modificar las propuestas –total o parcialmente- o si es positivo continuar como se previó, etc.

Allal, Cardinet y Perrenoud (1979, 132), refiriéndose a la evaluación formativa aplicada a los procesos de enseñanza y aprendizaje del alumno, señalan tres características que la distinguen.

- “La recogida de datos concernientes al proceso y las dificultades de aprendizaje encontradas por los alumnos.
- La interpretación de esta información desde una perspectiva Criterial, y en la medida de lo posible, diagnóstica de los factores que originan las dificultades de aprendizaje observadas en el alumno.
- La adaptación de las actividades de enseñanza y aprendizaje en función de la interpretación realizada de los datos claros recogidos”

En cualquier caso, si se consigue desarrollar la evaluación formativa con todas sus virtualidades, se evidenciará su funcionamiento como estrategia de mejora y se comprobará la mejor y mayor consecución de los objetivos propuestos. Es decir. Que no sólo mejorarán los procesos, sino que, en consecuencia, mejorarán tan bien palpablemente los resultados de aprendizaje que se alcancen por parte del alumnado. Si esto no ocurriera, no nos engañemos, algo estaría fallando. No pueden separarse procesos y rendimientos, porque constituyen dos caras de una misma moneda-. Siempre el resultado es consecuencia de un proceso. La persona es una y si los medios que utiliza para aprender son más adecuados a su personalidad y, por lo tanto, su proceso de aprendizaje mejora, también mejoran sus resultados o aprendizajes finales. La educación es un todo en el que, sólo a efectos metodológicos, desglosamos sus componentes. En la realidad educativa diaria cualquier ingrediente positivo que incorporemos favorecerá el conjunto y cualquier elemento negativo que aparezca distorsionará la actuación prevista. Es una observación que hay que tener en cuenta a lo largo de estas páginas, para no perder de vista la globalidad de la acción docente y la unidad de la persona-discente- a quien se forma y con quien se trabaja.

La evaluación según su normotipo

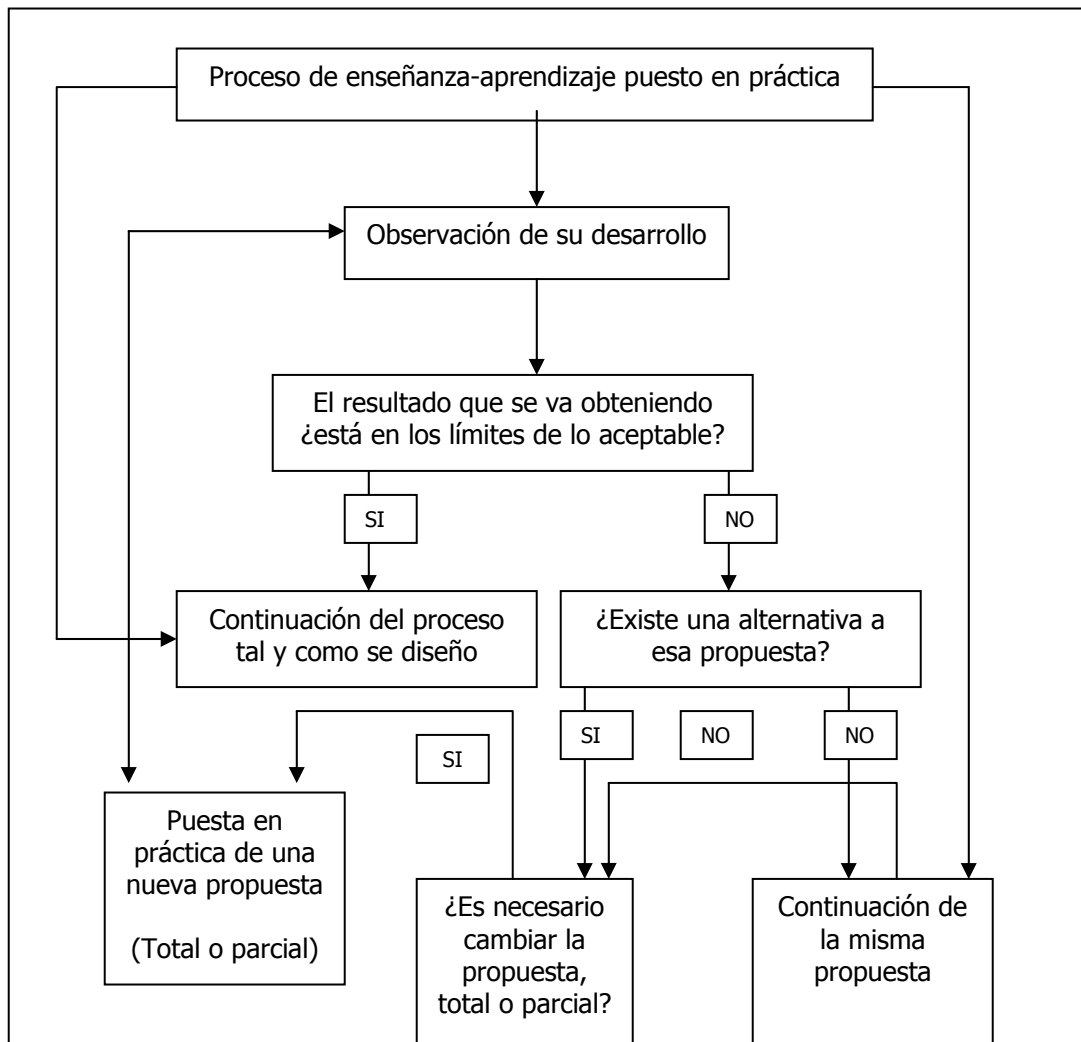
El normotipo es el referente que tomamos para evaluar un objeto/sujeto. Según este referente sea externo o interno al sujeto- en nuestro caso concreto de evaluación de aprendizajes en el alumnado-, la evaluación se denomina *nomotética* o *idiográfica*, respectivamente.

Evaluación nomotética

Dentro de la evaluación nomotética podemos distinguir dos tipos de referentes externos, que nos llevan a considerar la evolución normativa y la evolución Criterial.

La *evaluación normativa* supone la valoración de un sujeto en función del nivel del grupo en el que se halla integrado. Es decir, que si el nivel de los alumnos de un grupo es elevado, un alumno con un nivel medio puede resultar evaluado negativamente o, al menos, por debajo de lo que lo sería si estuviera en un grupo de nivel general más bajo. A la inversa, un alumno de tipo medio resulta evaluado de forma altamente positiva en un grupo donde el nivel general es bajo, cuando esta valoración no responde con exactitud a sus posibilidades reales frente a los referentes externos marcados por el sistema educativo.

Figura 8
Aplicación de la evaluación formativa
(Adaptación de Peterson, P. L., y Clark, C. M.: 1978)



Aunque el profesor inmerso en este tipo de evaluación no sea consciente de su forma de valorar, la descrita es una situación absolutamente generalizada. En efecto, hay que reconocer que es muy difícil -prácticamente imposible- mantenerse al margen del clima generado en un grupo por su nivel de aprendizaje y que se valora por encima de sus potencialidades reales al alumno que demuestra dominar algo más que el resto de los objetivos previstos en una programación.

Esta situación se demuestra objetivamente mediante la aplicación de pruebas estándares a una muestra significativa de alumnos y alumnas, obteniendo su valoración con unos criterios fijados de antemano y contrastando los resultados con las "calificaciones" anteriores dadas por el profesor a esos alumnos. Un ejemplo, experimentado por mi parte en 1983, fue la aplicación de las pruebas elaboradas por el ministerio de educación y ciencia para evaluar las enseñanzas mínimas del ciclo inicial de la educación general básica. Fueron seleccionados los centros donde debía realizarse la aplicación y, entre otros, resultó elegido uno público (Colegio A) de un barrio marginal, con la mayoría de su población en paro y una clase sociocultural nítidamente baja, a la vez, también apareció en la muestra un segundo (Colegio B) del centro de la ciudad, privado subvencionado – en aquel momento- y cuyos alumnos pertenecían claramente a una clase sociocultural medio/ alta. Los datos obtenidos de la aplicación aludida se volcaron en unos registros diseñados homogéneamente, en los que aparecía una comuna con la valoración obtenida por cada niño en a prueba estándar y otra paralela donde se reflejaba la valoración que se había dado al niño al final del curso anterior. En una comprobación superficial, seleccioné varios niños y niñas que habían sido valorados con "5" en la prueba estándar, perteneciente a los dos colegios, y, mientras en el Colegio A eran alumnos "sobresalientes" por sus calificaciones anteriores, en el Colegio B eran solamente "aprobados" o, incluso en el caso de una niña, estaba reprobada y permanecía en el ciclo (figura 9).

Otra situación muy frecuente es la que presenta a todo profesor cuando corrige los ejercicios o pruebas realizados por los estudiantes de un mismo grupo. Cuando comienza, de forma estricta, y un alumno no ha respondido a los objetivos que se pretendían y no domina un mínimo de contenidos previstos, le reprueba. El segundo, tercero, cuarto..., corren la misma suerte. Cuando ese mismo profesor está leyendo el ejercicio duodécimo, y encuentra dos cuestiones medianamente explicadas, dan un respiro porque ya alguien se ha enterado de algo, y aprueba a ese alumno. A partir de ahí, los criterios (¿) se van ampliando y flexibilizando y es probable que el ejercicio trigésimo apruebe con el mismo contenido con el que ha reprobado el primero. Pongo el interrogante al lado de criterios porque, efectivamente, de los que está adoleciendo esta evaluación es de falta de criterios claros, único, prefijados..., que sirvan a de referente para analizar el contenido del ejercicio y, por lo tanto, para establecer unas normas mediante "justas" en la valoración de todo el grupo.

En otros apartados de esta obra seguiremos comentando situaciones similares que aparecen en el modelo de evolución que actualmente se practica en nuestro sistema de enseñanza y que, urgentemente, deben ser superadas para remediar algunos –demasiados- efectos negativos que el sistema ejerce sobre las generaciones que lo padecen.

Es válida la evolución normativa cuando se pretende determinar la posición ordinal de un sujeto dentro de un grupo, en cuyo caso las normas de valoración estarán en función directa de lo que el conjunto del alumnado (u otro tipo de población) domina o deja dominar.

Pero la ordenación de los estudiantes v como “primero” o “último” del salón no me parece, precisamente, ni recomendable ni educativo, por lo que considero que ese tipo de evolución no resulta apropiada para nuestros fines, ya que carece de referencia válida y fiable tanto para la persona que se forma como para el profesorado que la valora.

La *evaluación Criterial*, precisamente, intenta corregir el fallo que plantea la evaluación normativa, y propone la fijación de unos criterios externos, bien formulados, concretos, claros..., para proceder a evaluar un aprendizaje tomando como punto de referencia el criterio marcado y/o las fases en que éste se haya podido desglosar. Fue propuesta por Popham, J. W. (1980), a la vista de las disfunciones que, permanentemente, se producían en la valoración del alumnado por la influencia, ya descrita, del nivel general del grupo en la valoración de cada uno de sus miembros.

En palabras de Popham, “una prueba que hace referencia a un criterio es la que se emplea para averiguar la situación de un individuo con respecto a un campo de conducta bien definido” (1980, 147-148).

El mismo autor señala que lo fundamental en la evolución Criterial se basa en:

- “1) La delimitación de un campo de conductas bien explicitado.
- 2) La determinación de la actuación del individuo en relación con ese campo” (1980, 151).

Para desarrollar una estructura correcta y valorar de este modo, en primer lugar hay que distinguir bien los objetivos de los criterios de evolución. Y., después, transformar sucesionalmente el objetivo en comportamientos observables y valorables (Orden, A, de la: 1982), que traduzcan lo que el objetivo pretende que la persona alcance y que resulte posible de evaluar mediante los criterios establecidos.

Los objetivos (más o menos generales, de una etapa, ciclo o curso, de un área curricular) marcan la meta a la que se pretende llegar. Los criterios³ de evolución deben especificar detalladamente cuando se considera que un alumno ha alcanzado un objetivo determinado. Un ejemplo: el objetivo didáctico de un ciclo de Educación Primaria, en el área de Lengua y Literatura, puede ser que el alumnado domine la ortografía. ¿Cuándo voy a afirmar que el alumno ha alcanzado ese objetivo? Aquí cabe especificar el criterio oportuno. Si es para el primer ciclo de la etapa, puedo considerar que un alumno ha logrado el aprendizaje si en sus escritos no comete más de un 10% de faltas de ortografía, de acuerdo con la secuenciación correspondiente del objetivo para este ciclo (saber utilizar el punto y aparte, el punto y seguido y la coma; usar correctamente las mayúsculas y aplicar la norma de escribir m antes de p y b).

La especificación detallada de estos criterios sirve, evidentemente para valorar de forma homogénea a todo el alumnado y para determinar el grado de dominio que debe tener un alumno en relación con el objetivo que se pretende.

No obstante, se corre el riesgo –ya experimentado en años anteriores por nuestra educación- de desglosar hasta tal punto las conductas o comportamientos que los estudiantes deben seguir (Bloom, B. S. y cols.: 1972), que tanto ellos como el profesorado queden atados de pies y manos en la programación de los procesos de enseñanza y aprendizaje y en su desarrollo, lo cual supone seguir un modelo absolutamente conductista del aprendizaje y su desarrollo, lo cual supone seguir un modelo absolutamente conductista del aprendizaje y es un peligro evidente que debe evitarse. Hay que señalar unos criterios de evaluación, pero con la suficiente amplitud que permita la práctica de un proceso realmente educativo, capaz de ser regulado en su cambio en función de las necesidades y circunstancias de vida persona y del entorno en el que tienen lugar la educación. Por lo tanto, teniendo en cuenta las aportaciones actuales de la psicología evolutiva y de la psicología del aprendizaje y las implicaciones que deben tener para las interacciones positivas en el aula, el criterio de evaluación, desde un punto de vista pedagógico, será la pauta o norma mediante la cual se lleve a cabo la evaluación tanto de los procesos como del logro de los objetivos que se hayan marcado, por áreas, para cada ciclo o curso de las diferentes etapas educativas. De esa forma, aunque esté marcado un nivel concreto de consecución, el profesorado lo podrá contextualizar automáticamente al ritmo y estilo de aprendizaje de cada sujeto y acompañarlo en función de sus peculiaridades. Ello implica poder modificar, con cierta flexibilidad, durante el proceso los planteamientos iniciales cuando se considere imprescindible, lo cual no obsta, como digno, para que en cada objetivo se identifiquen los criterios de evaluación más apropiados, que sean válidos como

³ La palabra "criterio" procede del griego. Juzgar, y es definida por el Diccionario de la Real Academia Española (1992) como: "Norma para conocer la verdad.// 2. juicio o discernimiento." Por su parte, ésta es la primera aceptación que el mismo Diccionario ofrece de "discernir": "Distinguir una cosa de otra, señalando las diferencias que hay entre ellas."

referentes equilibrados y homogéneos en la emisión de juicios de valor para la valoración formativa y estimuladora del alumnado.

Evolución idiográfica

Cuando el referente evaluador son las capacidades que el alumno posee y sus posibilidades de desarrollo en función de sus circunstancias particulares, es decir un referente absolutamente interno a la propia persona evaluada, la evaluación se denomina idiográfica.

El realizarla supone la valoración psicopedagógica inicial de esas capacidades del alumno o alumna, y la estimación de los aprendizajes que puede alcanzar a lo largo de un periodo de tiempo determinado (un curso, un ciclo). De acuerdo con esa valoración y estimación realizadas, el alumno va siendo evaluado durante su proceso e, igualmente, se valora el rendimiento final alcanzado. Si éstos coinciden, al menos, con lo estimado, se considera el rendimiento satisfactorio, de lo contrario., el rendimiento sería insatisfactorio, si bien puede estar ocasionado por diversas causas, incluso ajenas a la voluntad del alumno evaluado.

Este tipo de evaluación es positivo individualmente porque se centra totalmente en cada sujeto y valora, sobre todo, su esfuerzo, la voluntad que pone en aprender y formarse. Evalúa, en síntesis, lo más importante en la educación personal. Las actitudes. Pero choca con los planteamientos sociales de todo sistema educativo en los países del mundo desarrollando (capitalistas y competitivo por definición). Un alumno, así, puede haber alcanzado todos los objetivos que razonablemente debí tras un proceso de aprendizaje y, simultáneamente, estos objetivos pueden no ser suficientes al compararlos con los exigidos por el sistema educativo para otorgar una titulación, por ejemplo, o para promocionar a otro ciclo o curso. Con lo cual, estaríamos ante la situación de un alumno con valoración idiográfica satisfactoria, pero insuficiente ante los criterios externos marcados por el sistema general de enseñanza. También se nos presenta el caso contrario. El alumno que alcanza suficientemente los criterios establecidos desde el exterior, pero cuyo rendimiento es n es satisfactorio porque no cubre sus posibilidades de desarrollo, estimadas como más amplias.

Ambas situaciones resultan problemáticas, aunque quizá más la primera que la segunda. La segunda resulta más irritante e injusta frente a la anterior, porque, casi siempre, en estos casos se está dando preferencia a la consecución de objetivos que implican el dominio de contenidos conceptuales, frente a los procedimientos o actitudes, que son más educativos y más relevantes para el futuro desarrollo personal, estudiantil o profesional del alumno evaluado. Por lo que se refiere ala primera, cuando el alumno presente necesidades especiales de educación – de lo contrario no debería aparecer el caso descrito- o dificultades específicas provenientes de circunstancias particulares, hay pocas soluciones para

la "concesión" del título correspondiente –hacer un regalo de esta índole supone una estafa, tanto para el alumno como para la sociedad-, pero sí las hay para su orientación personal y profesional. Se debe buscar si integración más positiva en el entorno en que viva, mediante una profesión adecuada que le permita salir adelante siempre que será posible- y a través de la cual se realice y se sienta feliz. El título, en estos casos, suele importar más a la familia que al alumno, y no debe significar nada para que tenga lugar el desarrollo máximo de las capacidades del adolescente evaluado.

La conjunción de ambas evaluaciones –Criterial e idiográfica- debe aportar soluciones a los problemas expuestos. ¿Cómo? Posiblemente incorporando elementos ideográficos a los criterios de evaluación para el alumnado. Dando más importancia a los procesos de desarrollo actitudinal, social, efectivo, etc., que a la adquisición de aprendizajes puramente conceptuales. Adaptando a cada uno de los criterios generales de evolución establecidos: no eliminando lo imprescindible, pero si matizando o incorporándolo que resulte necesario para la formación de un estudiante concreto y personalizado. Esta labor ésta en manos de los educadores, cada día, cuando establecen la comunicación habitual con sus alumnos, esa comunicación educativa que es el eje de de todo proceso de enseñanza y que debe contextualizarse según la interacción particular que tienen en cada aula.

La evaluación según su temporalización

De acuerdo con los momentos en que se aplique la evolución, ésta puede ser *inicial, procesual o final*.

Evaluación inicial

La evaluación es aquella que se aplica al comienzo de un proceso evaluador, en nuestro caso referimos a la enseñanza y aprendizaje. De esta forma se detecta la situación de partida de los sujetos que posteriormente van a seguir su formulación y. por lo tanto, otros procesos de evaluación adecuados a los diversos momentos por los que pasen.

Tal situación de partida puede presentarse:

- a) Cuando un alumno llega por primera vez a un centro, bien para comenzar su escolaridad, bien para continuar. En el primer caso, será necesario realizar una amplia captura de datos para precisar del mejor modo las características de todo tipo del alumno (personales, familiares, sociales, etc.). Esta primera evaluación tienen una función eminentemente diagnóstica, pues servirá para conocer a ese alumno y poder adaptar al máximo, desde le primer momento, la actuación del profesor del centro a sus peculiaridades. En el segundo caso, es de suponer que el alumno aporte

su expediente escolar, de manera que la evaluación inicial estaría en función de los datos ya poseído y de los que falten para completar los necesarios en el nuevo centro. Para cualquiera de las dos situaciones se cuenta, habitualmente, con registros diversos —oficiales y particulares de cada centro— en los que anotar los datos recogidos (Knapp, R.H.: 1983).

- b) Cuando se comienza un proceso de aprendizaje concreto, como puede ser el trabajo con una unidad didáctica. En esta situación la evaluación inicial resultará útil para detectar las ideas previas que el alumnado posee en relación con el tema que se va a trabajar. Igualmente, se pondrán de manifiesto las actitudes hacia la temática —en su caso— y el mayor o menor dominio de los procedimientos que van a ser necesario para su desarrollo. Los medios para conseguir estos datos son variados: un coloquio, un debate, la realización de unos trabajos preparatorios..., pueden ser idóneos para su obtención. A partir de la información conseguida, se adaptará convenientemente el principio de la unidad didáctica programa, para adecuarla a los conocimientos en generales del grupo. Mediante el desarrollo oportuno se intentará que todos alcancemos los objetivos básicos e imprescindibles para poder seguir adelante en el proceso de aprendizajes subsiguientes. En algunos casos, incluso, habrá que posponer una unidad concreta, o, si no es determinante en ese curso ciclo se deberá aplazarse hasta otro curso posterior en el que los alumnos tengan la madurez o los conocimientos suficientes para poder asimilar los nuevos contenidos y alcanzar los objetivos propuestos en ella.

Sin una evaluación inicial falta el conocimiento previo que es preciso poseer de una persona en pleno desarrollo para poder adecuar la enseñanza a sus condiciones de aprendizaje, y cumplir de esta forma la función reguladora que hemos asignado a la evaluación.

Evaluación procesual

La evolución procesual es aquella que consiste en la valoración continua del aprendizaje del alumnado y de la enseñanza de profesor, mediante la obtención sistemática de datos, análisis de los mismos y toma de decisiones oportuna mientras tienen lugar el propio proceso. El plazo de tiempo en el que se realizará estará marcado por los objetivos que hayamos señalado para esta evaluación. Puede referirse al tiempo que dura el desarrollo de una unidad didáctica (una semana, quince días, 6/8 horas...), a un periodo trimestral de aprendizaje, anual, bianual (un ciclo completo de dos años)... Estará en función del tipo de aprendizajes que se evalúen (las actitudes es preciso valorarlas en plazos amplios, lo contrario es imposible y, por lo tanto, absolutamente superficial) y, además, hay que tener en cuenta que, desde esta perspectiva, se superpondrá unas y

otras evaluaciones procesuales: la de objetivos que implican asunción de actitudes a largo de un curso, con la de objetivos que suponen la adquisición de conceptos o del dominio de procedimientos, durante quince días o un mes. Ambas evaluaciones procesuales, cuyos resultados permanentes se van anotando en los registros preparados para ello.

La evolución procesual es la netamente formativa, pues al favorecer la toma continua de datos, permite la adopción de decisiones "sobre la marcha", que es lo que más interesa al docente para no dilatar en el tiempo la resolución de las dificultades presentadas por sus alumnos. Aparece un "error" que no usará para sancionar ni para calificar negativamente, sino que será útil para detectar el problema de aprendizaje que ha puesto de manifiesto; se resuelve mediante la adecuación de unas determinadas actividades o las explicaciones oportunas y se continúa el proceso de aprendizaje. Con este modo de actuar, será más fácil que la mayoría de los alumnos y alumnas lleguen a alcanzar los objetivos básicos propuestos para todos. Si el profesorado no posee estos datos hasta que ha transcurrido un mes, la oportunidad de subsanar las dificultades presentadas, que así es relativamente sencillo brindar a los alumnos, se hace prácticamente imposible, cuando se ofrece, ya suele ser tarde: el alumno ha perdido el interés o no puede combinar varios aprendizajes simultáneos para continuar al mismo ritmo del grupo.

Del mismo modo, llevar a cabo rigurosamente la evaluación procesual es lo único que permite mejorar el proceso de enseñanza, pues es durante el tiempo en que tienen lugar cuando se puede comprobar los fallos y los elementos que están funcionando positivamente, para –si es posible- subsanarlos o reforzarlos, respectivamente, de inmediato y, en un momento posterior, confirmar o reformular las líneas de programación con las que se trabaja.

La evolución final es aquella que se realiza al terminar un proceso –en nuestro caso, de enseñanza y aprendizaje-, aunque éste sea parcial. Una evolución final puede estar referida al final de un ciclo, curso o etapa educativa, pero también al término del desarrollo de una unidad didáctica o del proceso habido a lo largo de un trimestre. En definitiva, supone un momento de reflexión en torno a lo alcanzado después de un plazo establecido para llevar a cabo determinadas actividades y aprendizajes.

Es una evaluación en la que se comprueban los resultados obtenidos, aunque es necesario advertir que no por ello debe tener funcionalidad sumativa. Si coincide con una situación en la que tienen que decidirse definitivamente acerca de la obtención de un título, por ejemplo, será final y sumativa, pero si se sitúa al terminar el trabajo con una unidad didáctica, resultará simplemente final..., e inicial del trabajo que se va a realizar al día siguiente. Por ello, la evaluación final puede adoptar las dos funciones descritas anteriormente para la evaluación

adoptara las dos funciones descritas anteriormente para la evaluación: formativa y sumativa. Servirá, así, en su función formativa, bien para continuar adecuando la enseñanza al modo de aprendizaje del alumno., bien para retroalimentar la programación del profesor, quien, a la vista de lo conseguido, tomará las decisiones oportunas para mejorar el proceso de enseñanza en la unidad siguiente. En su función sumativa, resultará imprescindible para tomar la decisión última sobre el grado de lo alcanzado por un alumno y obrar en consecuencia. Los resultados de la evolución final, por otra parte, puede analizarse e interpretarse con tres referentes distintos:

- a) En relación con los objetivos y los criterios de evaluación establecidos para la unidad didáctica, el final del trimestre, curso o ciclo, de manera que se determine la situación de cada alumno en relación con los aprendizajes que institucionalmente se encuentran establecidos para ser conseguidos por todo estudiante al que haya que dar un título que avale su superación. Se tratará en este caso, por lo tanto, de una evaluación *nomotética Criterial*.
- b) En relación con la evolución inicial realizada a cada alumno y las posibilidades de desarrollo y aprendizaje que se estimaron podría alcanzar. Se determinará así lo satisfactorio o insatisfactorio de su rendimiento y se estará realizando, en consecuencia, una *evolución idiográfica*.

En relación con los resultados alcanzados por el resto del grupo, incluso, del conjunto del grupo en comparación con otros grupos de alumnos del mismo curso o ciclo en el centro o de distintos centros. En este caso se estará llevando a cabo una *evaluación nomotética normativa* que, si bien creo que no debería llegar al alumno o grupo particular, si puede resultar interesante e ilustrativa para el profesorado y para el centro, pues le sirve de referente para conocer su situación y valorar la calidad educativa que está ofreciendo a su comunidad.

Por último, sólo quiero hacer la observación de que evaluación final no tienen por qué coincidir con "examen", "control"..., o prueba semejante. Si el profesor considera que no posee datos suficientes para valorar lo conseguido por sus alumnos, puede, realizar un trabajo o prueba donde lo constante fehacientemente. Por si dispone de bastante información como para valorar con seguridad los aprendizajes alcanzados, la evaluación final constituirá un análisis y reflexión sobre los datos obtenidos, que llevarán a la formulación de la valoración correspondiente. Téngase en cuenta, además, que algunos de los objetivos educativos nunca van a poder comprobarse en una situación puntual como es la del tradicional examen: los relacionados con el desarrollo social, actitudinal, afectivo..., lo relativos a comunicación oral..., deben evaluarse procesualmente, y, como resultado de esta valoración continua, emitir el resultado final acerca del rendimiento logrado por el alumno.

La evolución según sus agentes

De acuerdo con las personas que en cada caso realizan la evolución, se dan procesos *autoevaluación, Coevaluación y Heteroevaluación.*

Autoevaluación

La Autoevaluación se produce cuando el sujeto evaluará sus propias actuaciones. Por tanto, el agente de la evaluación y su objeto se identifican. Es un tipo de evaluación que toda persona realiza de forma permanente a lo largo de su vida, una que continuamente se toman decisiones en función de la valoración positiva o negativa de una actuación específica, una relación tenida, un trabajo llevado a cabo, etc.

Al tratar a la Autoevaluación en el terreno profesional, hay que considerar la conveniencia, primeramente, de introducir su práctica de modo habitual entre los alumnos y alumnas. Con diferentes grados de complejidad, según las edades a las que nos referimos, el alumnado es perfectamente capaz de valorar su propia labor y el grado de satisfacción que le produce. Simplemente hay que darle pautas para que lo haga con seriedad y con corrección –no arbitrariamente ni por juego-. Y que sepa la influencia que su juicio va a tener en la valoración global que se realice posteriormente sobre su actuación y progreso. Al comenzar el desarrollo de una unidad didáctica, se facilitará a los alumnos la información detallada acerca de los aspectos que deben autoevaluar, para que puedan autoobservarse y examinar su trabajo continua y, así, llegar a conclusiones rigurosas al final del proceso, esta orientación del docente es importante –imprescindible- también para evitar la excesiva influencia de la subjetividad en la propia evaluación, mucho más frecuente cuando más inmadura es la persona que la realiza. La época de estudiante se caracteriza por ser esencialmente evolutiva y ambiente, con gran predominio de la subjetividad, dadas las características de inmadurez a que aludimos. Así, los alumnos con una autoestima baja, un temperamento depresivo, de tendencia pesimista o un medio familiar poco estimulante, infravalorarán sus trabajos, mientras que los optimistas, con alta autoestima o un medio familiar y social que los ayude, valorarán en exceso todo lo que realicen. Estas tendencias hay que encauzarlas y, precisamente, en este proceso de ajuste y equilibrio consiste ese “aprender a valorar” que se pretende en la educación.

En segundo lugar, muchos de los objetillos educativos de la educación obligatoria implican que el alumno sea capaz de “valorar” en términos genéricos, que luego se aplica diferentes cuestiones, según las áreas curriculares donde se incardinan. El patrimonio artístico y cultural, la riqueza lingüística, las actitudes solidarias o no discriminatorias etc. Para aprender a valorar, el único cambio existente es practicar valoraciones en distintas circunstancias y en relación con diferentes ámbitos. Una forma de evaluación es la autoevaluación del propio trabajo y la propia actividad.

Por eso, en este caso la Autoevaluación se convierte en un procedimiento metodológico para alcanzar uno de los objetivos educativos previstos. El de que el alumno sea capaz de valorar.

En tercer lugar, desde el enfoque de evaluación de la enseñanza o la práctica docente, la autoevaluación continua que realice el profesor de sus actividad en el aula y en el centro constituye un elemento imprescindible para mejorar paulatinamente los procesos educativos, ya que sin esa reflexión fallan los datos básicos para tomar decisiones correctas y oportunas. Muchas veces se deciden cambios sin fundamento..., y todos sabemos que cambia por cambiar no conduce a nada: se acierta o se falla por puro azar, minimamente orientado en función de la experiencia que tenga el profesor en cuestión. Las opciones en uno u otro sentido deben estar asentadas en reflexiones evaluadoras rigurosas: la adopción o rechazo de una metodología, de un tipo de actividades, de un programa..., será el resultado de una valoración de la teoría y práctica anteriores con ellos.

Coevaluación

La *Coevaluación* consiste en la evaluación mutua, conjunta, de una actividad o un trabajo determinado realizando entre varios. En este caso, tras la práctica de una serie de actividades o al finalizar una unidad didáctica, alumnos y profesores pueden evaluar ciertos aspectos que resulten interesantes destacar. Tras un trabajo en equipos, cada uno valora lo que le ha parecido más interesante de los otros, por ejemplo. En un coloquio, se valora conjuntamente el interés de las actividades, el contenido de los trabajos, los objetivos alcanzados, la suficiencia de los recursos, actuaciones especialmente destacadas de algunos alumnos, etc. Es posible, igualmente, pasar un cuestionamiento –anónimo- a los alumnos, para que opinen con absoluta independencia sobre lo realizado, y contrastar así con lo percibido por el profesor o profesora.

Son diferentes los aminos mediante lo cuales llevar a cabo la Coevaluación, pero es importante tener en cuenta que, si no hay costumbre en el grupo de realizarse prácticas de este tipo, debe comenzarse por valorar exclusivamente lo positivo. Las deficiencias o dificultades surgidas las valorará el profesor. No se debe olvidar que los alumnos tienen la misma visión o percepción de “para qué se evalúa” que muchos profesores y que la sociedad en general, y que habitualmente esta valoración se realiza para resaltar lo negativo, lo, mal hecho, para sancionar, para “reprobar”. Y eso es lo que hacen ellos en cuanto tienen oportunidad: decir lo que han hecho mal los otros, con lo cual los efectos de la Coevaluación pueden convertirse en la disgregación del grupo y el rechazo de todos contra todos. Hay experiencia muy negativa de esta práctica y, por ello, hay que tomar todas las precauciones necesarias. ¿Es interesante? Sí. Sin duda alguna, pero se dan dos situaciones claras, con fases intermedias entre una y otra:

- a) Si el grupo de alumnos viene realizándola habitualmente, deberá poseer una visión positiva de la evolución; evaluar constituirá una fase del proceso educativo que sirve para mejorar poco a poco el propio aprendizaje y todo cuando ocurre en el aula, por lo cual es favorable y beneficioso para el grupo.
- b) Si el grupo de alumnos nunca la ha realizado, habrá que comenzar por explicar cuál es la finalidad de la evolución y, más en concreto, de la Coevaluación. Y no sólo hay que explicarlo; hay que demostrarlo con la práctica habitual del profesor en el aula, ahí es donde el grupo se convencerá de las virtualidades de la evolución, de nada vale que el profesor "diga" que los fines de la evolución se centran en la mejora de todos, si luego examina y la evaluación sólo sirve para aprobar o reprobar. El conjunto de prácticas evaluadoras debe responder a un mismo planteamiento, coherente con la teoría que se intercambie con los alumnos. En caso contrario, es mejor no introducir nuevas prácticas que pueden redundar negativamente el proceso educativo. En esta segunda situación que planteamos es donde resulta preferible comenzar evaluando aspectos positivos del trabajo, de manera que no se dé pie a que surja alguna actitud negativa que deteriore la práctica comenzada.

Heteroevaluación

La *Heteroevaluación* consiste en la evaluación que realiza una persona sobre otras: su trabajo, su actuación, su rendimiento, etc. Es la evaluación que habitualmente lleva a cabo el profesor con los alumnos, y a cuyo proceso se dirigen principalmente las páginas de esta obra.

Es un proceso importante dentro de la enseñanza-como y ha quedado puesto de manifiesto- rico, por los datos y posibilidades que ofrece y complejo por las dificultades que supone el enjuiciar las actuaciones de otras personas, más aún cuando ésta se encuentra en momentos evolutivos delicados en los que un juicio equivoco, "injusto", poco sopesado., puede crear actitudes de rechazo (hacia el estudio, hacia la sociedad) en ese niño, adolescente o joven que se educa.

Consideraciones finales

Para terminar este apartado quiero insistir en algo ya apuntado antes: la tipología de la evaluación expuesta es válida para presentar claramente las virtudes que ofrece la evolución desde diferentes enfoques o perspectivas, pero en la práctica, en su aplicación al quehacer diario pueden darse mezclados algunos de los tipos descritos y, sobre todo, se conjuga su funcionalidad para conseguir la meta común: la mejor formación del alumnado. Así, tanto la evaluación inicial como la final pueden ser, a la vez, formativas o sumativas- según la funcionalidad con la que se apliquen-; lo mismo ocurre con la autoevaluación o Coevaluación, que,

además, pueden ser iniciales, procesuales o finales. La evaluación Criterial o la idiográfica se aplicarán para perfeccionar o para seleccionar, al principio, durante o al final de un proceso e incorporando a ellas autoevaluaciones o coevaluaciones del propio alumnado... la acción educativa es variada y requiere soluciones diversificadas que permitirán que la enseñanza de un profesor y el aprendizaje de un alumno caminen al unísono. Este aprendizaje, en definitiva, es un continuo que no se rompe ni se divide en compartimentos estancados. Por ello, la acción educativa eficaz debe plantearse también como tal y, en consecuencia, la evolución no puede aplicarse aisladamente ni sin interferencias permanentes de todas las perspectivas. La evaluación final es la evaluación inicial del día siguiente...

La ordenación de la labor entre docente y discente se realizará seleccionando el tipo de evaluación más adecuado para cada situación y los objetivos que se pretendan con ella, eligiendo técnicas e instrumentos para llevarla a cabo y poniéndola en práctica, colegiadamente, del mejor modo posible. Pero una vez organizado el proceso –de forma rigurosa pero flexible- en camino se presentarán multitud de ocasiones en las que será necesario echar mano de todo el tipo de recursos para resolver determinadas circunstancias que se planteen con ciertos alumnos: de ahí la conveniencia –la exigencia profesional- de disponer de cuantos conocimientos evaluadores-en este caso- sea posible y aplicarlos en los momentos del camino en que resulten útiles para superar una situación.

La calidad de la evaluación depende, en buena medida, de la rigurosidad de la evaluación. Y evaluar no siempre resulta fácil, pero siempre es ineludible. La evaluación es compleja porque desemboca no sólo en asumir, sino en rechazar muchas posibilidades, y la abundancia de posibilidades implica una drástica y, a veces, dolorosa selección. Por eso, quizá, resulte más fácil rechazar con certeza que asumir sin dudas. Por eso, una vez más. Insistimos en las ventajas de valorar no sólo lo negativo (lo claramente rechazable), sino todo lo que de positivo aparezca en el proceso educativo: analiza las ventajas de cada elemento ofrecerá criterios válidos en los que apoyar la elección de una ruta y no la de otra similar.

Un automovilista, al evaluar una situación y realizar cierta maniobra. Puede jugarse la vida; un médico, al valorar el estado de un paciente ponerle un tratamiento, puede poner en peligro su salud, lo mismo que un juez la libertad de un acusado o un profesor al valorar a un alumno puede –por “indulgente” o “exigente” –poner en peligro su adecuación. Estos riesgos, ¿no son lo suficientemente graves como para ser minimizados mediante la asunción de las necesarias precauciones con el fin de brindar las mayores ganancias a los afectados? La libertad individual que todo profesional debe tener para desarrollar su trabajo – la libertad de cátedra, por ejemplo- no puede implicar el ejercicio de la arbitrariedad. Por eso, si la evaluación debe ser objetiva, hay que saber que será más objetiva más colegiada, ya que la

subjetividad de la persona se objetiva al contrastarse con otras subjetividades y regir los juicios por criterios equivalentes- y previamente establecidos- para todos.

En enseñar puede ser fácil. Lo difícil es enseñar bien. Enjuiciar a alguien puede hacerse apresuradamente, pero evaluar a una persona o un proceso educativo completo es una tarea seria y compleja. El denominado "fracaso escolar" esconde un número de personas concretas que se han malogrado desde el punto de vista educativo. ¿Qué agricultor se sentiría satisfecho si todos los años se le malograra el 30, 40 ó 50% de su cosecha? ¿Qué fábrica podría permitirse el lujo de desechar el 30, 40 ó 50% de su producción? Ya ha quedado dicho: descalificar o reprobar es fácil, reorientar, reenseñar para que alguien aprenda lo que antes no aprendió excluye ser un docente repetitivo- casi mecánico-; exige ser flexible e innovador, creativo, evaluador, conocedor de que muchos y distintos caminos conducen a la misma Roma. El conocimiento, las destrezas, las actitudes no son algo que se posee o no se posee, sino un hacerse, desarrollarse, conformarse permanentemente, aunque también un olvidarse, perderse, bloquearse o inhibirse.

El sujeto que aprende no es una computadora a la que se debe instalar un programa, sino un ser inteligente que se debe formar para que pueda ser libre y asuma ser responsable. Justamente para ser libres y responsables es preciso poder y saber evaluar, de manera que algo que también tienen que enseñar el docente a salud alumnos como antes quedó señalado- es a evaluar y evaluarse, por eso la autoevaluación y Coevaluación son básicas y constitutivas de una educación que pretende procesos libres, auténticos, comprometidos e innovadores. Conocer, juzgar, valorar..., son palabras sucesivas en un proceso educativo en el que el sujeto madurara equilibradamente para saber, saber hacer y saber ser con los otros.

OBJETIVOS DE LA EVALUACIÓN

De acuerdo con lo expuesto hasta ahora, los objetivos que permite perseguir la aplicación de un modelo evaluar cualitativo, formativo y continuo, adaptado convenientemente a los procesos de enseñanza y aprendizaje, no se limitan ya a comprobar lo aprendido por el alumnado cada cierto tiempo, sino que amplían sus expectativas y posibilidades y cubren un campo más extenso, más completo, pues la evaluación se incorpora, desde el principio, al camino del aprender y enseñar, y ofrece, en consecuencia, mayores aportaciones y apoyos al conjunto del procesos que transcurre.

En concreto, los objetivos que pretende la evaluación, considerando el concepto adoptado para la misma y las diferentes perspectivas comentadas en su tipología, serán:

1. Detectar la situación de partida general para dar comienzo a un proceso de enseñanza y aprendizaje
2. Facilitar la elaboración de la programación idónea y adecuada para los alumnos y alumnas, en función del diagnóstico realizado en el paso anterior.
3. Durante la aplicación de cada unidad didáctica.
 - a) Conocer las ideas previas del alumnado.
 - b) Adaptar el conjunto de elementos de la unidad a la situación del grupo.
 - c) Regular el proceso de enseñanza y aprendizaje: reforzando los elementos positivos, eliminando los elementos negativos, adaptando las actividades a las posibilidades de cada alumno, superando de inmediato las dificultades surgidas.
 - d) Controlar los resultados obtenidos.
 - e) Mantener los objetivos no alcanzados, incorporándolos a unidades siguientes.
4. Confirmar o reformular la a programación en función de los datos obtenidos con el desarrollo de las unidades didácticas que la componen.
5. Orientar al alumnado para futuros estudios o salidas profesionales.
6. Elaborar informes descriptivos acerca del proceso de aprendizaje que sigue cada uno de los alumnos.
7. regular y mejorar la organización y actuación docente, tanto en su perspectiva con respecto al centro como para su actividad en el aula.
8. Controlar el rendimiento general del alumnado, para su oportuna promoción o titulación.
9. Seleccionar los recursos didácticos y programas específicos para el centro.

El disponer de permanente información sobre el proceso educativo que está teniendo lugar y también, acerca de los resultados últimos conseguidos o de la idoneidad mayor o menor de ciertos programas o recursos, permite claramente conseguir los objetivos aquí señalados.

**RUEDA, Beltrán Mario, Díaz Barriga Arceo Frida,
“Consideraciones teóricas sobre la evaluación en la educación”
En: Evaluación de la docencia. México Ed. Paidós pp 23-36.**

CAPÍTULO I

**CONSIDERACIONES TEÓRICAS SOBRE LA EVALUACIÓN EN LA
EVALUACIÓN***

Jacques Ardoino
(Universidad de París VIII)

El propósito de este trabajo es recordar que la evaluación, inseparable del control- y sin embargo distintos entre sí, heterogéneos-, supone una problemática, antes de considerar la como un instrumento. El control de los conocimientos, de la calidad de un producto, de las exigencias de cierta actividad, de la ejecución de un programa, de la puesta en marcha de una política, del cumplimiento de una tarea, constituye una función crítica simple, necesaria, que pretende verificar la conformidad en relación con una norma, un modelo o programa, o por lo menos la medida en que se aleja de éste. Esto se realiza a través de un conjunto de procedimientos que suponen una homogeneidad entre la norma, el modelo, el programa y los fenómenos que se reportan, para compararlos. Con este solo hecho se eliminan las particularidades, principalmente las temporales.

Por el contrario, la evaluación es un proceso temporal que requiere para un mismo objeto múltiples cuestionamientos, generalmente heterogéneos, si no es que contradictorios. En ese sentido, el control es siempre monorreferencial, mientras que la evaluación es explícitamente multirreferencial e Intersubjetiva, en la medida en que pasa obligatoriamente por testimonios y presentaciones, estas dos funciones críticas son indispensables; sin embargo, no corresponden a los mismos valores (conformidad y coherencia, por una parte; sentido y pertinencia, por otra): finalmente deben articularse y conjugarse, mas no confundirse, como sucede comúnmente, no pueden realizarse al mismo tiempo mediante un solo procedimiento. Sus particularidades respectivas son *paradigmáticas* y *epistemológicas*, más que *metodológicas*.

EL DESARROLLO DEL PROYECTO Y SU PROCESO DE EVALUACIÓN

La sola intención educativa requiere hablar de un proyecto [Ardoino 1987]. Precisemos este punto. La educación es el ejercicio de una función social, aunque parezca limitada a las dimensiones de una acción personal. En efecto, provienen generalmente de un sujeto más viejo, más experimentado, profesional o familiar, más competente, o de una persona moral encargada de dicha función, frente a un (os) sujeto (s) más joven (es), menos informados (s). en un sentido mucho más

general, amplia que la sociedad reconoce la necesidad, para su propia sobrevivencia y desarrollo, de preparar y formar a sus integrantes para cumplir lo mejor posible las «condiciones» (humana, profesional, social...) de su aculturación (respecto a las leyes y los valores establecidos, adquisición del saber-hacer y de las competencias).

A diferencia de la pedagogía, que de acuerdo con su raíz etimológica¹ se constituye en servidora (puesto que es más pragmática e instrumental en su diseño de optimización del hecho de transferir conocimientos),² la educación hace referencia a los valores, asume como misión proponer para compartir (no sólo «saber» y «saber-hacer», sino también «saber ser») [Ardoino 1963], cuando no para inculcar. En el centro del proyecto no son solamente los «contenidos», sino las opciones más fundamentales las que se cuestionan. ¿Cuáles son las habilidades que se desea transmitir? ¿Qué modelo de sociedad se persigue? ¿Qué tipo de alumnos, estudiantes, hombres, se desea formar mediante dichos procedimientos? Más concretamente: qué «representaciones» se forman el educador, profesor e instructor de sus compañeros, de sus aptitudes para resolver los problemas que enfrentan, de sus capacidades de creación y de invención, de su «negatividad» (elaboración de contraestrategias en respuesta a las estrategias de las cuales se siente objeto), de su necesidad de *transgresión*, siempre contemporánea de la necesidad de aceptar una *ley* común. Por lo mismo, la educación, dado que es dialéctica, es propia y explícitamente filosófica, mientras que la pedagógica pretende ser científica (si no es que militante), para favorecer la permanencia del positivismo de fines del siglo pasado, en el marco de la «instrucción pública» republicana.

En Francia, la rehabilitación de la noción de preproyectos (originalmente más extendida en los medios confesionales) por los aparatos de Estado (educación nacional, ministerios Savary [1982] y Jospin [1990], después del largo debate público que se dio a partir del caso expuesto por la «bella provincia» canadiense al inicio de los años ochenta,³ consagra a su manera el regreso explícito de la filosofía a las ciencias humanas. Se termina por hablar, paradójicamente, de *Proyecto pedagógico* [Ardoino 1987, Boutinet 1990], en la medida en que resulta más juicioso optar por la idea de proyecto educativo. Sin embargo, la ambigüedad no desaparece. De igual modo - sin no es que en mayor medida - que las dimensiones ideológicas, políticas o filosóficas siempre presentes, el aspecto instrumental tenderá a imponerse. La noción de proyecto al igual que la de

¹ El pedagogo era en la antigüedad el esclavo encargado de llevar a la escuela a los niños de la casa a la que pertenecía. Se puede hablar aun hoy, en un sentido más amplio (finalidades asignadas por un mercado, orden social), de una función «doméstica» de la educación, obstruyendo su función eventualmente «política».

² Por su parte, la didáctica (de las disciplinas) se interesa más específicamente todavía en las condiciones de la transmisión de un saber a partir de un punto de vista, privilegiando la naturaleza misma de la disciplina. ¿Cómo enseñar las matemáticas en función de su idealidad?

³ Véase, en Québec, los «Livres» de J. Morin [1982] terminado en la «Ley 40r».

evaluación, sobre la que se hablará más adelante, se parece, analógicamente hablando, a una elipse con dos centros de *sentidos* relativamente heterogéneos, al menos teóricamente, el proyecto expresa, en primer lugar, una intención filosófica o política, una visión, el enunciado de una misión, la expresión de una afinidad, suponiendo de forma un poco indeterminada, indefinida, sin no es que infinita, un horizonte de valores representados en busca de realización o más, comúnmente, afirmando la existencia de sujetos en busca de tales valores. Es evidente, esta búsqueda no podría realizarse más que a través del tiempo, en el marco de un tiempo venidero, de un futuro no precisamente programable. Sastre [1960] hablada de un proyecto del hombre «haciéndose a través de lo que hace». La idea misma de un «proyecto de sociedad» proviene de esta aceptación. La referencia tanto a un imaginario social como a una imaginación radical reforzándose mutuamente [Castoriadis 1975] es, por supuesto, fundamental; manteniéndose en la misma perspectiva teórica, la traducción estratégica, operatoria, precisa, determinada, de esta visión, formulada a través de objetivos estrictamente definidos. Se trata esta vez de una dimensión esencialmente programática del mismo conjunto, donde el plano de una máquina concebida por un ingeniero, la maqueta de un inmueble o de una obra de arte elaborada por arquitecto o, más aún, el proyecto de presupuesto para un responsable financiero, pueden proporcionar buenos ejemplos. Sus realizaciones posteriores, si se dan, deben parecerse lo más posible a las anticipaciones.

Idealmente, el «propósito» y el «programa» son los dos constituyentes necesarios de esta dinámica, en cuyo seno lo imaginario y lo real deben estar claramente articulados. Cuando, en este conjunto, no hay una fuerte de sentido explícito y renovado, el protocolo es insignificante (éste puede ser el caso de la pedagogía por objetivos); cuando el objetivo no se concreta, la intención se debilita. Sin dicha articulación habría, por lo tanto en ambas partes, mutilación, dolencia. Asimismo, es necesario comprender que esa anticipación debe evitar retomar el método hipotético-deductivo tradicional, en un intento por proyectar en el porvenir aquello que se deduce del análisis del presente como bastante frecuente; esto se ha ensayado en numerosas previsiones, futurizaciones, gestiones de prospectiva, a riesgo de la esterilidad. Lugares, funciones, roles y efectos de la *sorpres*a y de lo *imprevisto*, el ardid propio de los efectos «inesperados» de determinada estrategia, no se dejarán encerrar fácilmente en la demasiado cómoda patología de los «efectos perversos», por lo que son extremadamente importantes. Son determinantes, tanto por la concepción, la conducción y la puesta en marcha del proyecto, como por la evaluación de cada una de sus diferentes fases.

Pero en los hechos abundan los proyectos programáticos, ajenos a sus fuentes de sentidos posibles (dicho de otra manera, «reificados»). Las «peticiones» masivas de proyectos a las instituciones (escolares, educativas sociales, psiquiátricas...) han obtenido siempre le mismo resultado. en cursos de las dos alternativas antes citadas,, el Ministerio de Educación Nacional ha amontado con la intermediación de

rektorados e inspecciones académicas, centenas de «proyectos vacíos», solicitados de acuerdo con la supuesta intención de la autoridad. Son, en efecto, proyectos que como documentos han sido transmitidos a la autoridad. Dicha autoridad parece haber cometido en este caso error de ver el proyecto, por una parte, como un instrumento movilización (casi mecánica) de las instituciones, de las personas y de los equipos y, por otra, como un medio para justificarse a sí misma, pregonar su proyecto arguyendo una intención seudodemocrática: «Cuando los proyectos d en las base sean conocidos, la cima deducirá más fácilmente el suyo.» La inconveniencia, siempre burocrática, en el nivel de las instancias locales, de esta exhortación autoritaria (llamada iniciación) a producir su «proyecto pedagógico» o su «proyecto de institución »⁴ reside principalmente en la ignorancia de que se trate de un proceso- no de un instrumento- que puede, después de un tiempo conveniente, permitir la dinámica deseada.

Se podrían contar otros ejemplo: un plan de desarrollo económico, financiero o industrial, o bien la apuesta en práctica de una política de alfabetización o, más Aún, un programa de educación sanitaria. En todos los casos, *las preguntas relativas al sentido* habrían implicado que los diferentes protagonistas, los actores- autores, se encontraban efectivamente asociados, en vista de esta evaluación, con las personad que aportan igualmente sus contribuciones en las forma de recursos metodológicos y de mi radas exteriores.

Mientras que la responsabilidad de control con frecuencia era confiada, a través de los apartados, de los dispositivos, a personal especializado, a veces suboficial, situando en un nivel un poco más elevado en la jerarquía, la evaluación implicada indistintamente a todos los protagonistas involucrados en la situación. Entonces, los evaluados son al mismo tiempo los evaluadores.

Esta autoevaluación no lo es de manea absoluta, porque, por la ausencia de toma de distancia y de puntos de vista exteriores, se correría el riesgo de desembocar en posiciones exageradamente sesgadas, esto quiere decir que no se podría hablar propiamente de evaluadores pensando a la manera del control; se trata de profesionales de la valuación especialmente formadas para ese efecto. Serán esencialmente recursos humanos que funcionen, más como consultores que como *expertos*¹⁰, que ayudarán a los grupos, a las colectividades, a las organizaciones, a escoger una metodología propia para la evaluación.

⁴ El proyecto de la institución, que supone igualmente un proyecto pedagógico sólo el que se apoya, se diferencia esencialmente por sus multidimensionalidad. Implica parte explícitante económica, financieras, organizacionales, de los participantes.

¹⁰ A diferencia del experto que aporta, de manera más puntual, las competencias y conocimientos, generalmente especializados, (de punta)(, para ponerlos a disposición de sus cliente afín de darle, en forma de soluciones, los medios para resolver los problemas que ha encontrado; el consultor efectúa más bien, dentro de una relación de acompañamiento, más durable, un trabajo que tiende a la evaluación de las representaciones, de las actitudes, a ver los comportamientos de sus cliente evolución de la cual se esperan justamente las mejoras destacadas. Vèase Jacques Ardoino 1990.

En la perspectiva que aquí se desarrolla, con lo cual se pretende teorizar acerca de las prácticas evaluativas, se trata por lo tanto de comprender que con el control y la evaluación (en el sentido pleno de este último término) se tienen dos acepciones diferentes de la defunción crítica, que cada vez ponen en juego sus paradigmas respectivos, eventualmente complementarios, a partir de una distinción previa. Estos dos procedimientos de ninguna manera se duplicarán, y sí se revelan uno y otro necesarios.

Pero, en contrapartida, el control está inmerso en la evaluación, en la medida en que ésta es considerada en su acepción más amplia, queriendo discriminar entre los datos constitutivos de una situación, entre las diferentes facetas presentadas por un objetivo, para así poder estimular a o apreciar su valor, preguntándose sobre los sentidos que la han conducido, en función de los diferentes contextos a que se refiere (y la intención como la capacidad de medida que la acompaña), en tanto *uno de los medios* de los que se vale la evaluación. No sucede igual en sentido inverso. El control –definido como investigación sistemática a través de un dispositivo apropiado, que pone en juego los procedimientos según una secuencia, lo más breve posible, con el propósito, con el propósito de establecer la conformada, si no es que la identidad, entre una norma, un modelo y los fenómenos que se comparan, en o en su defecto, de permitir la medida de su alejamiento- no contienen jamás en sí mismo a la evaluación, en tanto que interroga sobre el sentido,. Finalmente, lo que hace la diferencia es más la cantidad de la mirada puesta sobre el objeto que la naturaleza de ese último o los instrumentos empleados (“rejillas, escalas de medida).

El problema se plantea casi en los mismos términos cuando se trata de la investigación científica. Si ésta es concebida únicamente según un modelo canónico que privilegia la administración de la prueba, o a falta de formas de demostración que partan de razonamientos lógico-matemáticos, la evaluación de los procedimientos, de las experiencias, de los productos, se efectúa esencialmente en términos de control y de verificación. En el momento en que la investigación, principalmente en el campo de las ciencias del hombre, se concibe desde un punto de vista de un análisis completo de los efectos del sentido, más que de una explicación siempre posible de los efectos de fuerza, la evaluación – sin excluir el control y el análisis cuantitativo, como tantos otros medios- se amplía hasta suponer igualmente los enfoques alternativos, históricos, etnológicos, hermenéuticos, plurirreferenciales, reconociendo en los sucesivos, del solo hecho de tomar en cuenta la intencionalidad, un status a la implicación, a lo imaginario, a lo inconsciente, etc., a través del juego de las representaciones. Se está entonces en el marco de una intersubjetividad, más que en el de la objetividad.

Finalmente, respecto del polo más explícitamente praxiológico del eje donde se inscribe, la relación entre la evaluación y el proyecto es evidentemente estrecha. Cuando el proyecto no es pensado más que de manera programática, lo que desafortunadamente es el caso de la mayoría de los proyectos de instituciones, puede dar naturalmente material al control, pero también pueden suscitar de forma provechosa una evolución, en términos, de interrogarse sobre el sentido (a

falta del cual, aun en el caso de que los programas sean eficaces, nada les impediría a éstos tender a la insignificancia). En la medida en que un proyecto diseñado existe y es efectivamente reconocido en la formulación de sus "finalidades". Diferenciales de sus enunciados programáticos y estratégicos en términos de objetivos, no puede ser evaluado (en el sentido pleno del término) más que por aquellos que lo concibieron, asistidos o no por recursos humanos externos. Ello no excluye evidentemente al control. Pero, a su vez, el proyecto diseñado, cuando contienen enunciados explícitos, es también la fuente necesaria de sentido a la que uno se referirá cuando se trate de evaluar efectivamente, y no sólo de controlar las prácticas. También aquí la relación de un proceso con la temporalidad, por el carácter a la vez *pasado* (raíces, orígenes, hechos fundadores, memorias), *presente* (realización) y *futuro* del proyecto, esta bien marcada. La evaluación propiamente dicha queda ligada a la historia de un fenómeno, de una situación, de una persona, de una comunidad, etc.

Desde el punto de vista de un análisis más institucional, la paradoja de una puesta en marcha de un proyecto (pedagógico, educativo), a sí como del proceso de evaluación de las realizaciones a las que este ha permitido llegar, reside de probablemente en que sus diseños se apoyan siempre, al inicio, en una valoración de lo "local", que reclama una autonomía cuando menos relativa, para al final lograr casi siempre someterse al conjunto a un control jerárquico de nueva cuenta reforzado. Convenientemente manejada, la puesta en marcha y la evaluación – es decir, en tanto no permanezcan perdidas en el frenesí de una tecnicidad seca-pueden al menos contribuir al enriquecimiento esperado de un juego interactivo.

EL ENTRETEJIDO DEL CONTROL Y LA EVALUACIÓN DE LAS PRÁCTICA EDUCATIVAS

Como ya se afirmó, desde muy jóvenes y con la formación precoz de nuestra sensibilidad evaluamos espontáneamente aquellos que nos gusta o nos disgusta, aquellos que preferimos o aquellos que no. en este sentido, evaluar está ligado a la facultad de discernir, reconocer, diferenciar, distinguir, juzgar, apreciar, estimar, involucrarse, y somos progresivamente llevados a elaborar muy temprano, individual y colectivamente, *las prácticas evaluativas*. Pero en el sentido más técnico, ahora la evaluación de signa, principalmente, un dispositivo institucionalizado, a veces instrumentados, tan racional como sea posible, que tienen por objeto llenar las mismas funciones críticas pero de manera más sistemática, en una perspectiva sobre todo praxiológica (optimación de la acción, ayuda a la decisión), lo que no excluye que la evaluación pueda incluso contribuir, llegando el caso, a una empresa más científica de generación de nuevos conocimientos. Son primeramente la pedagogía y la docimología, con sus objetivos de calificación de los comportamientos, y después las ciencias económicas, tomando en cuenta su actual tendencia a una relativa hegemonía, las que han contribuido al desarrollo y la multiplicación de tales dispositivos, y conducen al mismo tiempo a intenta precisar, si no es que a teorizar su metodología.

Hoy la evaluación se encuentra íntimamente asociada a las tareas de gestión (dirección, *management*), a las grandes orientaciones políticas, a su administración (evaluación de las políticas públicas), a la administración (evaluación de las universidades, de las escuelas, de los diplomas, de los profesores de los estudiantes y los alumnos). Se puede querer muy legítimamente evaluar a las personas y a los productos, los proyectos y sus resultados, los procesos y sus procedimientos, las políticas y sus programas.

Se encuentra uno así en presencia de una gran diversidad si no es que de una dispersión de desordenada; en otras palabras, de una gran heterogeneidad. En efecto, desde ahora el término *evaluación* no parece cubrir, de maneja bastante confusa, dos subconjuntos que, si estuvieran convenientemente distinguirlos, presentarían, cada uno en su seno, una mejor cohesión e incluso mayor coherencia. Por un lado, los *procedimientos de control*; por el otro, la *evaluación* propiamente dicha, en el sentido de *proceso* y gestión [Ardonino 1976]. El control, apunta, en favor de tales procedimientos, a establecer la conformidad o identidad entre una norma y los fenómenos que se reportan para fines de la comparación. Esto es una verificación.

Cuando el control no se quiere esencialmente «en tiempo real» (informatizado), está fuera del tiempo, tanto por razones políticas (igualdad, no arbitrario, universalidad) como técnicas (eficacia). Supone correspondencia, homogeneidad, entre el modelo –norma, anterior y exterior a la operación de control propiamente dicha, y los elementos que se comparan. Su cuestionamiento es único (monorreferencial). Las afinidades con la medida, lo cuantitativo, son indiscutibles. La ficción de la objetividad pura le conviene para las consecuencias cercanas de la relatividad (tríptico indisociable. Observatorio-observador- cosa observada). En tanto que proceso- por lo tanto evidentemente más cualitativa-, la evaluación es esencialmente temporal, inscrita en una duración, y por este hecho inacabada e interminable, corresponde mejor entonces a una gestión que a la forma cerrada de una secuencia o de una sesión. Más allá del juego de interacciones inteligentes y deseables que la constituyen, su orden/ desorden es ante todo el de la ínter subjetividad. Plurirreferencial en consecuencia, la evaluación lleva, a partir de sus múltiples perspectivas, a cuestionamientos sobre el sentido y las significaciones de los fenómenos a propósito de los cuales una se interroga, más que en relación con su conformidad hacia las normas. Mientras que el control permanece indiferente, la evaluación queda inseparable de una hermenéutica. Da materia a una interpretación [Ardoino y Berger 1986]. Así, uno puede interrogarse sobre las modalidades, las causas, los disfuncionamientos constatados dentro del sistema educativo, produciendo o agravando las desigualdades., generando el fracaso escolar a partir de preguntas clásicas: «¿Cómo?», «¿por qué?», que conducen a pregunta por la calidad de los estudiantes, de los profesores, o incluso sobre la naturaleza y amplitud de los medios a sus disposición. En relación con éstos, la pregunta por qué (por quién) abrirá nuevas pistas radicalmente diferentes (matriz eventual de una sociedad a dos velocidades y de productos, en consecuencia, diferencia, inscritos en el mismo aparato escolar).

En un buen número de situaciones, lo que comúnmente se llama evaluación no es más que control disfrazado. Controlar el paso de cada uno de los estudiantes en un programa de pedagogía por objetivos, así como las calificaciones y los niveles logrados, es una cosa; evaluar la pertinencia de ese programa considerando los objetivos perseguidos en función de los contenidos curriculares asimilados y su relación con el nivel de satisfacción, otra muy diferente, la evaluación tratará sobre objetos temporales, inscritos en una historia que cuestionará en primer lugar a partir de una pluralidad de valores heterogéneos., en función de su proyecto, la evaluación deducirá enseguida criterios, eventualmente *indicadores* apropiados, y cuando sea posible, en otros casos, serán los acontecimientos analizadores los que permitirán la evaluación.

¿ES POSIBLE LA EVALUACIÓN DE LAS ACTIVIDADES PEDAGÓGICAS DE LOS PROFESORES EN LA UNIVERSIDAD?

La evaluación de los profesores en una u otra universidad o en el conjunto de universidades de un país, es un caso particular de evaluación al cual pueden aplicarse los planteamientos generales formulados anteriormente. ¿Qué quiere decir evaluar a los profesores? Quiere decir evaluarlos desde el punto de vista de una competencia intrínseca y otros no. Pero pensar las cosas de esta manera pone en evidencia que esto es tan sólo una forma de entender «evaluar». Se evalúa a los profesores como un objeto que tienen alguna propiedad o algunas dimensiones de esta propiedad, tal o cual escala de medida. Evaluar a los profesores es evaluar las relaciones de enseñanza-aprendizaje, pero considerando la particularidad de los estudiantes, quienes, aun estando en grupos, nunca son uniformes y estarán cada uno con su historia, su competencia, sus dificultades, etc. La evaluación así concebida es otra cosa. En el primer caso se evalúa a las personas con cuestionarios, contando sus publicaciones o midiendo pretendiendo medir sus propiedades, sus características, en el segundo, se evalúan situaciones que son intersubjetivas y mucho más móviles, mucho más inciertas, en el primer caso, es comprensible que a los directivos de las universidades les gustaría evaluar de esta manera, porque se cuenta y se mide y es un proceso más general u abstracto. En el segundo caso, la evaluación no puede ser más que particular, ligada a esos estudiantes y a ese profesor específico, y no puede generalizarse. Se está por lo tanto en dos tipos de función crítica, las dos necesarias pero complementemente heterogéneas. Es conveniente que se hagan las dos, pero no hay que confundirlas, no se debe creer que evaluar a los profesores de la primera manera es aquello que se llama evaluación; es controlar. Se trata de controlar a partir de un modelo, de una norma supuesta, se dice que alguien está ubicado por abajo o por encima; él es mejor que otro, se les compara, exactamente como se ha clasificado a los alumnos durante siglos-. Clasificar a los alumnos quiere decir jerarquizarlos y compararlos de la misma manera, se hablaba en otra época del alumno capaz, inteligente, menos inteligente, con aptitudes, etc. Con lo anterior se quiere subrayar lo irracional de aceptar una «evaluación» de parte de las

universidades para la gestión del personal, de los recursos humanos y de los recursos para la enseñanza. Por el contrario, lo que es necesario evaluar son las relaciones, los pros y contras, las dificultades entre el profesor y los estudiantes, en este sentido, los mejores evaluadores sean los estudiantes, los mismos profesores y probablemente será necesario añadir una medida exterior para que no se encierren en sí mismos. Pero esta mirada exterior no es jamás una norma o un modelo que aporta la verdad. El evaluador no es tan profesional que tiene el poder o la competencia de evaluar; se trata de desarrollar un proceso que tienen una función crítica, compartida entre profesores y estudiantes, para plantearse una serie de preguntas. Pero preguntas y formularse un conjunto de preguntas es exactamente aquello que se dijo al inicio, al referirnos a lo multirreferencial. No hay una pregunta verdadera-falsa, más o menos en relación con una escala preestablecida. Hay cuestionamientos múltiples, contradictorios entre ellos, heterogéneos y que ayudan a que la función crítica aquí sí sea, cualitativa más que cuantitativamente, más rica y desarrollada.

BIBLIOGRAFÍA

- ARDOINO, Jacques, *Imaginaire dans la formation permanente*, París, 1976.
_____, *Propos actuels sur éducation*, 6a. ed., París : Gauthier Villars, 1978,
_____, «Projet éducatif, projet de société »*Pour*, no, 94, París, 1984.
_____, «Postures (on impostures) respectives du chercheur; de Pextect du consultant », *les nouvelles formes de la recherche en éducation an regard di une Europe en devenir*, Actas del coloquio inenacional de la AFIRSE, París ; Matrice-ANDSHA 1990.
_____, «Projet, formation-action, Education Permanente, no. 87, París, 1987,
_____, «Finalmente il n'est jamais pédagogie seans project», en Ardonio 1987.
_____, «Pédagogie de projet on protect éducatif», en Ardoino 1987.
_____, «Au filigrane d'un discours, la questions du contrôle et de l'évaluation» en el prefacio de Michel Morin, *L'imaginaire dans la formations permanente*, París, 1976.
_____, y G. BERGER, «L'évaluations au pouvoir» *Pour*, no. 107, París, 1984.
_____, y G. BERGER, *D'une évaluation en miettes a une évaluation en acte*, París Matrice-ANDSHA, 1989.
- BOUTINET, JEAN-PIERRE, *Anthopologie du pojet*, París : PUF, 1990.
- CASTOPRIADIS, CORNELIUS, *L'institution imaginaire de la société*, París Seuil, 1975,
- FAYOL, HENERI, *Administrations industrielle et générale*, París : Dunod, 1956.
PEYREDITE, ALAIN, *Le mal francais*, París : Plon, 1976.
SARTRE, JEAN-PAUL., *Critique de la raison dialectique*, París : Gallimard, 1960.

LAFOURCADE, D. PEDRO. “La evaluación en el proceso educacional” En: Evaluación de los aprendizajes. Argentina Ed. Kapelusz. Pp. 15-25.

1

LA EVALUACIÓN EN EL PROCESO EDUCACIONAL

1.- ETAPA DE UN PROCESO SISTEMÁTICO

Para comprender en su más amplio sentido el significado que se dará al término evaluación en los capítulos que integran este manual, será conveniente analizar una serie de ideas previas vinculadas al complejo que da existencia a este proceso. La evaluación.

La educación es entendida aquí como un proceso sistemático destinado a lograr cambios duraderos y positivos en las conductas de los sujetos sometidos a su influencia, en base a objetivos definidos de modo concreto y preciso, social e individualmente aceptables, dignos de ser sufridos por los individuos en crecimiento y promovidos por los responsables de su formación.

De acuerdo con esta definición, sólo podría admitirse que alguien se ha educado en una determinada dimensión cuando ha logrado –después de un período de enseñanza- una urdimbre de conductas inexistentes en el momento de comenzar sus aprendizajes. El logro de dichos cambios representan, pues, una meta que la educación intentará alcanzar a través de toda su estrategia.

Tales modificaciones (un aumento en la comprensión del uso de los números, una mayor apreciación del arte en la vida del hombre, la transformación de una actitud prejuiciosa por otra objetiva y equilibrada, un mejor conocimiento de la vida en la edad moderna, etc.) representan metas declaradas con antelación a la puesta en juego de un sistema metodológico que intente su realización en los educandos. Como es de suponer, no siempre dicho sistema logra su cometido sea por factores dependientes de su propia estructura o de aplicación, o derivados de una ambiciosa fijación de objetivos de un desconocimiento de la capacidad de aprendizaje de los alumnos o de la incapacidad de gobierno del docente. ¿De qué manera saber, entonces, que a través del proceso educacional se han cumplido de modo total o parcial los objetivos que se habían seleccionado con vistas a dirigir intencionalmente la acción didáctica? ¿Cómo advertir la medida en que se han logrado en uno, algunos o todos los que intervinieron en el periodo de aprendizaje? ¿Cómo saber que el conocimiento verbal que revela un grupo es un índice de su capacidad para interpretar, analizar, aplicar u ordenar algo? Todos estos interrogantes encuentran respuestas a través de un programa de evaluación,

cuyo fin primordial será averiguar en qué medida se han logrado los cambios de conducta previstos en los objetivos.

En los párrafos anteriores se han determinado en forma implícita diversas etapas del proceso educacional: fijaron de metas, ordenamientos de la estrategia metodológica, aprendizajes, evaluación y reajuste. No se podría pensar en la realización completa del acto educativo si faltara cualquiera de dichas etapas. Sin la indicación de objetivos, el proceso sería un barco a la deriva, sin un buen complejo metodológico, una acción insegura u azarosa, sin aprendizajes, un esfuerzo inútil; sin evaluación, una empresa de la cual se desconocería su eficiencia, pese a que la tuviera, y si8n reajuste, una tarea medias.

Siendo la etapa de un proceso, es lógico suponer que adquirirá sentido en el contexto total del mismo. Si la escuela otorga prioridad al campo de la información factual, la evaluación se reducirá a la elaboración de ítems que den cuenta de la cantidad de información acumulada. Si la estrategia metodológica ha reforzado el aprendizaje del espíritu crítico o de la amplitud mental, la evaluación consistirá en un conjunto de estímulos de elaboración más compela que intentarán verificar la existencia de dichas conductas y la medida aproximada de su desarrollo. Si el alumno no aprendió la técnica de inferir conclusiones o predecir cursos de acción, una prueba que contenga elementos estimulantes de dichas conductas, casi con seguridad no será resuelta satisfactoriamente.

Si la educación es un proceso sistemático, las etapas que lo integran también estarán influidas por esta característica.

Esta afirmación parece obvia. Sin embargo en la realidad no siempre se completa el proceso o se loase a medidas, presumiendo que el cumplimiento de un conjunto de actos didácticos, supuestamente acertados por el docente, debe originar necesariamente resultados positivos. Sólo d se puede estar seguro de afirmar la eficacia del esfuerzo docente si se aportan pruebas fehacientes de dicha eficacia: si, en suma, se produce a la evaluación de los resultados mediante técnicas confiables y válidas.

En resumen, la evaluación es entendida aquí como una etapa del proceso educacional que tienen por fin comprobar de modo sistemático en qué medida se han logrado los resultados previstos en los objetivos que se hubieran especificado con antelación.

2.- LA ANTIGUA CONFUSIÓN ENTRE MEDIR Y EVALUAR

La definición anterior ya está expresando, de modo implícito, que el concepto de *evaluación* es más amplio que el de *medición*. Véase el siguiente ejemplo: Alicia sube a la balanza de una farmacia y advierte que con sus 30 años y 1,72 m de altura, pesa 57kg. Baja radiante de satisfacción y orgullosa de la cifra registrada.

José también registra su peso: 83kg con 1,70 m y 39 años. No baja muy contento ni9 parece orgulloso del dato obtenido. Cualquiera que hubiera observado a la pareja y no comprendiera el significado de las cifras señaladas por la aguja, tampoco entendería el porqué de sus distintas reacciones. .

Con lo cual se llega a una primera conclusión; los guarismos que indican el resultado de una medición no parecen tener sentido por si mismos.

¿Qué advirtió Alicia en la lectura de su peso que la puso tan contenta? Pues que de acuerdo con las tablas de peso, altura y edad inscriptas en la propia balanza, se hallaba encuadrada dentro de la normalidad más rigurosa.

¿Y José? Lamentablemente se había excedido en por lo menos 13 kilogramos, de acuerdo con la misma tabla, y eso no estaba bien. Para lograr la misma satisfacción de Alicia cuando se pesara, debería adelgazar por lo menos esos kilos de más.

En ambos casos, los protagonistas han efectuado una evaluación de las mediciones observadas. Las han analizado, y asignado un sentido en relación a un patrón existente que les ha indicado cuáles sin los kilos -termino medio- que deberían pesar para considerarse normales. Cualquier alejamiento de la norma hacia arriba o hacia abajo, representará medidas de la anormalidad.

En síntesis, la evaluación es una interpretación de una medida (o medidas) en relación a una norma ya establecida.

Cabe una pregunta: ¡Sólo se aplicará esta definición a las mediciones cuantitativas, tal como la que se ha indicado en el ejemplo? Acaso cuando se observa a diario que Ricardo jamás reiteradas negativas de Pablo a prestar ayuda a sus compañeros, ¿No se concluye que carece de sentido de cooperación? Y en todos estos casos, ¿no se está interpretando un conjunto de hechos en relación a algún patrón, tal vez subjetivamente existente, pero patrón al fin? En efecto, tanto las mediciones cuantitativas como las descripciones cualitativas, sometidas a una interpretación y concluidas en un juicio de valor, constituyen aspectos de la evaluación.

Con los reparos que alguien quisiera oponer, es a fin de cuentas justificable afirmar que cuando se cuentan los ítems acertados de una prueba de rendimiento y frente a los no acertados se asigna un puntaje, se está efectuando una medición cuantitativa. Si ese puntaje se aleja de la medida del grupo, o se acerca a ella, en un sentido o en otro, se lo interpretará como muy bueno o menos bueno.

Del mismo modo, si a través de registros anecdóticos se verifica que un grupo de alumnos en quince3 ocasiones distintas ha demostrado que un grupo de alumnos en quince ocasiones distintas ha demostrado marcada responsabilidad en situaciones en las cuales pudieron no haberla tenido, se llegará a admitir la existencia de dicho tipo de haberla tenido, se llegará a admitir la existencia de

dicho tipo de conducta,. Y se concluirá que efectivamente son responsables. las múltiples observaciones (descripciones cualitativas) mediante una interpretación dada (comparación con los patrones escritos o no que determinan, por la característica de las acciones y la reiteración de las conductas, cuando se puede hablar de que alguien es responsable) han generado un juicio de valor (ya que ser responsable es mejor que no serlo).

3.- ¿QUÉ UTILIDAD SE PUEDE EXTRAER DE LA EVALUACIÓN?

El señor R. es maestro de séptimo grado y la señorita E. es profesora de historia de una escuela secundaria. Ambos han tomado exámenes en sus correspondientes ámbitos escolares. Han leído, corregido y devuelto los trabajos con las calificaciones de rigor, por rara casualidad la mayor parte de sus alumnos han sido aplazados-. Al entregar las pruebas, ninguno de los docentes expresó más comentarios que algunas generalidades sobre la flojedad de los trabajos,,la necesidad de mayor estudio, la promesa de levantar la calificación en proporción al esfuerzo que empeñaran..., pero nada más. A renglón seguido anunciaron los nuevos temas que tratarían durante los próximos días. Terminada la hora, también por casualidad, se encontraron en la sala de personal. Comentando sus respectivas experiencias en relación a las pruebas escritas en cuestión, intercambiaron opiniones y concluyeron que el tendal de aplazados era consecuencia de la típica falta de dedicación al estudio de los alumnos de esta época, que ambos eran sumamente exigentes en la asignación de notas y que ello se refleja en los resultados obtenidos y otras consideraciones que no vienen al caso comentar. Pero lo importante de lo acontecido es que a ninguno de los dos se les ocurrió pensar qué decisión tomar respecto de los objetivos que, según el examen, no fueron alcanzados por los alumnos, ni tampoco qué cambios introducir en la estrategia didáctica para no repetir lo sucedido, o reflexionar que si no habían aprendido el tramo del programa ya enseñado, no se podrán esperar milagros en el aprendizaje de los próximos contenidos, íntimamente ligados a los anteriores.

Por lo visto, el único objetivo del examen promovido por ambos docentes ha sido el de otorgar una nota. Lamentablemente esta visión simplista, precaria y estrecha de la evolución, no tienen cabida en una organización escolar que pretenda ofrecer un buen servicio educacional.

Un profesional de la enseñanza sabrá extraer múltiples utilidades de los resultados de un programa de evaluación aplicado a su grupo escolar. Sin entrar a considerar el aprovechamiento que pueda significar para el psicólogo escolar, el orientador, los padres o la administración de la escuela, la evaluación construye una actividad que permitirá al docente.

- a) Saber cuales objetos fueron cumplidos a través del ciclo didáctico proyectado.*

La posibilidad de logro de los objetivos concretos que el docente selecciona como tarea previa a la enseñanza, no constituye más que una hipótesis que sólo será válida con la confrontación de los resultados obtenidos.

En una unidad de aprendizaje, bien se pudo suponer que objetivos tales como ser capaz de:

1. Llevar a cabo cinco experimentos sencillos de tipo que aparecen en la sección 22 de libro de ciencias;
2. explicar las relaciones entre las plantas, los animales y su medio ambiente;
3. llevar a cabo simples disecciones de plantas y animales;
4. escribir un breve artículo sobre la evolución de las plantas y los animales, etc.

Serían cumplidos inexorablemente en el plazo de diez días, para lo cual se monto el complejo metodológico que debería realizarlos.

De presumir a concretar efectivamente hay un trecho que no siempre es corto. Ausencias de alumnos, feriados, mayor extensión de ciertas experiencias, planteamientos ambiciosos y demás factores bien pudieron haber efectuado el cumplimiento de alguno de los cuatro objetivos citados. Esta información sólo puede proporcionar, de modo preciso la evaluación que se realice acerca de la cuatro metas previstas. Los resultados revelarán que, por ejemplo, muy pocos alumnos comprendieron las relaciones existentes entre las plantas, los animales y el medio ambiente o que nadie sabe llevar a cabo un sencillo experimento del tipo que sugiere el objetivo.

Sin un programa que compruebe la validez de los esfuerzos realizados, dicha efectividad podrá quedar en el terreno de las conjeturas.

- b) *Intentar un análisis de las causas que pudieron haber motivado deficiencias en el logro de las metas propuestas.*

Los resultados que se deriven de un proceso de evaluación representarán un estímulo para que el docente busque explicación a las deficiencias observadas. Pasará revista a las principales causas, y las analizará con agudo sentido crítico. Los objetivos específicos de la unidad, ¿fueron previstos en función de las posibilidades de aprendizaje del curso? ¿Los alumnos fueron motivados suficientemente como para mantener un ritmo de interés uniforme a lo largo de la etapa de estudio? ¿No se habrá abusado de la exposición verbal? ¿Se habrá distribuido racionalmente el tiempo? ¿No se habrá pasado con demasiada rapidez la etapa del repaso y reajuste de lo aprendido? Las experiencias organizadas, ¿habrán sido las que más convenían? ¿Se debe suponer que durante los diez o doce días que duró la unidad los alumnos (y el docente) ha perdido

irremisiblemente el tiempo? ¿No será que el propio examen está mal construido y no mide lo que realmente debería medir?

La propia responsabilidad moral de quien debe dar cuenta de su eficiencia contribuirá a que se detecten con total objetividad, la o las causas que provocaron deficiencias en los rendimientos generales y específicos de la unidad. Esta etapa dentro del proceso de la evaluación es absolutamente necesaria e imprescindible para cualquier acción ulterior.

c) Adoptar una decisión en relación a la causal que concurrió al logro parcial de los objetivos revistos.

De nada serviría saber el porqué de la precariedad del rendimiento si no se intenta remediar la situación.

Si los objetivos son inadecuados al nivel de aprendizaje del grupo habrá que trasladarlos a otro grado o curso. Si los contenidos son demasiado difíciles deberá procederse en el mismo surtido. Si las deficiencias son subsanables dentro de la misma situación, no habrá otra salida que volver a enseñar lo que no fue aprendido., no se podrá adoptar la posición cómoda de continuar con otros temas cuando la mayoría de los escolares desconoce el anterior. Si los alumnos superaron realmente los objetivos, pero la prueba, debido a su deficiente construcción, no produjo la información correcta, la línea de ajuste, como es obvio estará apuntada en otros sentidos. Saber que un grupo pequeño o grande no superó importancia objetivos de un tramo del programa e inculpar dicho resultado al poco interés de los alumnos, a su crónica falta de aplicación al estudio, al "mal de la época", etc., y continuar el desarrollo de la asignatura como si nada malo hubiera ocurrido, no parece ser la mejor salida del problema.

d) Aprender de la experiencia y no incurrir, en el futuro, en los mismos errores.

Si se ha descubierto que los métodos adoptados no han favorecido sólidos aprendizajes, sería muy poco inteligente persistir en el uso de los mismos.

Si los exámenes revelan que el grupo carece de la experiencia básica y de la disposición necesaria para enfrenta4 futuros compromisos, sería poco acertado insistir en la misma dirección.

La comprobación de los resultados de los aprendizajes lleva implícita una evaluación de todos los factores que contribuyeron a su realización. Los puntajes medios y las desviaciones Standard de las pruebas, duran la medida de la eficacia de los instrumentos y de la organización impuesta a la empresa educacional y, al mismo tiempo, las bases para orientar una política de ajuste a ser emprendida a partir de la enseñanza del próximo sector de programa que corresponda.

Desde este punto de vista la evaluación contribuye a la constante reelaboración de la estrategia docente e impide la fijación de pautas rígidas e inamovibles en la conducción del proceso educativo.

Parece común insistir con el resto de posibles utilidades que se desprenden de una análisis prolijo de la comprobación del rendimiento: favorecer el agrupamiento de alumnos, promoverlos de grado, fase, curso, o nivel; ayudar a descubrir aptitudes, etc. Otros manuales las desarrollan con suficiente extensión y los docentes han aprendido a aprovecharlas desde antigua data.

En el desarrollo de este punto sólo se ha querido llamar la atención sobre aquellas cuestiones tomadas en mente en el quehacer didáctico, pero que existe una importancia capital para el correcto funcionamiento de la enseñanza.

4.Lo viejo y lo nuevo del problema

Analizando la antigua y la nueva bibliografía sobre el tema se advierten, más que espectaculares innovaciones, enfoques mejores consolidados por las conclusiones de variedad de investigación de base estadística y por concepciones más científicas de campo pedagógico.

El proceso evidenciado en la construcción de pruebas elaboradas por los propios maestros y profesores ha derivado, de algún modo, del mejoramiento paulatino introducidos por los expertos en la confección de pruebas estandarizadas de rendimiento. La inventiva en la idealización de ítems estructurados sobre variedad de bases y aplicables a propósitos muy definidos corresponde, en la mayoría de los casos, a los especialistas en mediciones y evaluaciones escolares. Con las reservas que suelen imponer apresuradas generalizaciones, bien se podría admitir que el progreso de las técnicas e instrumentos de evaluación no ha sido una consecuencia directa del esfuerzo docente, sino de los que se han dedicado de lleno a estudiar el problema y a resolver puntos espinosos muy controvertidos en el pasado. ¿Qué es lo que todavía se mantienen de los antiguos conceptos sobre evaluación? Resulta un poco difícil efectuar un deslinde entre lo nuevo. Quizá lo más importante que se haya conservado sea; el lugar que el cabe dentro del proceso educacional, la idea de que las apreciaciones globales emitidas subjetivamente por los docentes carecen, por lo común, de la suficiente validez y confiabilidad para ser tenida en cuenta, y la tremenda experiencia lograda sobre los errores y aciertos de los que iniciaron y desarrollaron el movimiento de la comprobación de los resultados del aprendizaje sobre bases más racionales.

¿Cuáles son las ideas que han ganado terreno en este campo?. En análisis de trabajos actuales ha permitido seleccionar las más evidentes:

- a) Salvo excepciones ya no se afirma que las pruebas “objetivas” son mejores que las de composición o ensayo, o que hay que eliminar los exámenes orales porque son reliquias de la edad de piedra de la educación. Se ha tratado de proveerles un sistema de procedimientos y técnicas de aplicación que los torne más eficientes y aptos para lo que intente medir. Ésta es la novedad. Según la característica de los objetivos supuestamente cumplidos y listos para evaluar, así también será la elección que se efectúe.
- b) El constructor de una prueba ya no elaborada ítems en forma indiscriminada. Subgrupos de elementos responden únicamente a un tipo de conducta que se pretende estimar o provocar. La antigua elaboración no preveía esta circunstancia y, por lo general, sólo media la cantidad de conocimientos acumulado en una o dos semanas de aprendizaje. En la actualidad, tal como se verá en los próximos capítulos en algunos ítems de interpretación el conocimiento está a la vista del alumno. El deberá extraer inferencias o determinar si los datos son suficientes para explicar una conclusión o señalar los errores existentes. Intentar medir, además del número de conocimientos que el examinado posee, el uso funcional que realiza de los mismos.
- c) Si el aprendizaje se logra cuando una conducta llega a ser algo distinta de lo que era antes de la efectiva influencia de aquel proceso, y si el la evaluación consiste en averiguar la medida en que esas conductas constituyen un patrimonio, real del sujeto, se forma necesario efectuar un recuento de las conductas posibles de lograrse en el plano escolar. Como las mismas son objetivos posibles de la empresa educacional, expertos e instituciones han trabajado empeñosamente para clasificarlos y ayudar a maestros y profesores a emplearlos con acierto.
Entre esas clasificaciones ya han adquirido amplia generalización las de Benjamin S. Bloom y colaboradores, Leo Nedelsky, Roberto L. Ebel, Nolan C. Kearney y las del servicio de Mediciones Educativas de Princeton por citar algunas de las más conocidas. Los docentes, a diferencia de sus colegas de épocas anteriores, tienen a su disposición largas listas clasificadas de objetos concretos, lo que les facilita enormemente la tarea de elaborar sus planes didácticos y el sistema de pruebas que debe comprobar las metas propuestas para su enseñanza.
- d) La representabilidad de las conductas incluidas en una prueba constituye otro aporte digno de mencionarse. Si existen cinco objetivos que se deben medir, y todos tienen el mismo grado de importancia, tal vez sea lo más

legítimo suponer que si el examen constara de veinte ítems, habría que elaborar por lo menos cuatro para cada objetivo.

- e) La mayor flexibilidad en la elección de los materiales que se emplean para construir un ítem es una conquista de los tiempos que corren. En una prueba no sólo intervienen textos escritos. Se ha difundido el uso de bases no verbales las cuales no requieren lectura comprensiva (un factor que a veces afecta la validez de un ítem), tales como dibujos, esquemas, gráficos, fotografías, mapas, tiras didácticas, material real, modelos, etc. Ello permite ampliar las posibilidades de las mediciones y evaluaciones a variedad de situaciones de prueba mucho más diversas que las tradiciones.
- f) Los clásicos ítems individuales, que sólo median conocimientos aislados, se han complementado con las que en el capítulo 2 se las ha identificado como pruebas multi-ítem de base común, las cuales consisten en un ejercicio introductorio y un grupo de ítems que suelen implicar el uso de distintas operaciones mentales y en donde el alumno debe resolver desde diversos ángulos una situación que por lo común es muy compleja. Precisamente han sido ideadas para evaluar resultados del aprendizaje del más alto nivel.
- g) Las pruebas prácticas, de ejecución funcionales, han adquirido un procedimiento asombroso en los programas de evaluación de esta década. No se desconocía su importancia, pero su uso está relegado por las pruebas de lápiz y papel. Múltiples proyectos de mejoramiento de la enseñanza de las ciencias e idioma, artes, artes plásticas y estudios sociales, han revitalizado el valor de estas pruebas, otorgándoles las técnicas necesarias para aplicarlas. Es evidente que su validez es mucho mayor que las de lápiz y papel, por ejemplo, y que la información que se puede lograr mediante su interpretación supera a la de cualquier tipo de examen. Actualmente integran los programas de evaluación de todos los niveles escolares.
- h) La preocupación más marcada de docentes e investigadores parece estar centrada en la búsqueda de instrumentos de medición válidos y confiables que estimulen el pensamiento creador (o divergente según la clasificación de Guilford). Las pruebas en las cuales el alumno debe seleccionar una respuesta no son adecuadas para medir la capacidad de formular un problema nuevo, organizan un contexto original o plantean hipótesis propias. Hasta el momento las de ensayo o las que describen las conductas en su situación real serían las más apoyadas por los que pretenden averiguar la existencia y medida de dichos objetivos. Dentro de muy poco tiempo se conocerán conclusiones dignas de crédito de los sistemas de pruebas ensayados en este campo menos investigado.

- i) Una evaluación integral de los objetivos que persigue la enseñanza no se agota en el área de los conocimientos y de las capacidades intelectuales. Parece ser que ninguna escuela organizada científicamente deja al azar la formación de actitudes (sociales, científicas), intereses vocacionales y personal, ajustes sociales y emocionales, hábitos de estudio, apreciaciones estéticas, etc. con el mismo celo que se monte el complejo metodológico para enseñar terminología y hechos específicos, principios u generalizaciones, teorías y estructuras, también se arquitectura lo necesario para enseñar aquellas conductas típicas del área efectiva.

Los instrumentos de evaluación habitualmente (pruebas orales, escritas objetiva, etc.) ya no serán eficaces para evaluar los objetivos concretos que también enuncian en dicha área (objetivos que tradicionalmente se declaraban de modo vago y general y no hallaban, como consecuencia, cabida en un programa de evaluación). Por tal razón se verifica, en esta aparte final de la década, un uso cada vez mayor de técnicas de observación, tales como los registros de vida escolar, las hojas de cotejo y las escalas de calificación, alcanzando también amplia difusión como recursos portadores de datos para un buen programa de evaluación, la entrevista, el cuestionario, el inventario y las técnicas sicométricas.

En síntesis, los antiguos exámenes y sus precarios resaltados, han dado paso a un sistema de comprobación mucho más enriquecido y técnicamente más preciso y confiable.

5. EL PAPEL DE LOS PROFESIONALES DE LA ENSEÑANZA Y DE LAS ORGANIZACIONES ESCOLARES EN EL PROBLEMA DE LA EVALUACIÓN

La nuevas generaciones que van obtenido sus habilitación docente como maestros o profesores poseen, por obra natural del progreso de las ciencias pedagógicas y áreas auxiliares requerirán mayor preparación que sus antepasados, tal como ocurrirá con los que les sucedan.

El análisis de los planes y programas de las escuelas normales y demás instituciones pedagógicas de nivel universitario, permite suponer que en los próximos años los elementos docentes del país habrán revitalizado, con un nuevo espíritu de cambio, la mayor parte de los esquemas conceptuales y de las realizaciones prácticas ya inútiles en los tiempos en que el hombre se está alistando para conocer personalmente su sistema planetario.

Lamentablemente, la acción individual de los miles de docentes que poblarán estas anchas latitudes se verá totalmente limitada so no concurren en su ayuda otros factores de peso decisivo en la posible transformación del sistema educativo. De nada valdrá que las instituciones formadoras hayan efectuado manifiestos

esfuerzos para proporcionar la mayor capacitación para los nuevos papeles y exigencias, si la realidad escolar que deben enfrentar persiste en conversar lo incoservable y en refugiarse en la comodidad de la rutina o en la complacencia de un conservadorísimo anacrónico. Inútiles serán los impulsos creadores del que ingrese como docente en una escuela donde el director haya creado un clima de obsecuente respecto por las reglamentaciones en vigor y la tradición de no hacer nada que rompa la armonía de lo que siempre se hizo y que nadie encontró malo porque así se hiciera. De más estarán las protestas del que llegue a otra escuela y observe que los alumnos son examinados del modo más precario y deficiente y calificados según los caprichos del que juzga, sin que ello signifique un problema para nadie y menos para el que dirige la escuela a o la supervisa.

Las digresiones anotadas tienen por fin advertir al lector que cualquier proceso de cambio en el ámbito educacional (contenidos más acordes con la época, métodos más eficaces, relaciones interpersonales más ajustadas a la función formadora del que educa, sistemas de evaluación más científicos, prácticos y realistas, etc.) debe contar con la acción coordinada de todos los que tienen relación con el mismo (directivos del organismo máximo que dirige la educación de una zona, región provincia o Estado, supervisores, directores de escuelas, personal técnico y docente, comunidad, etc.). si tal cosa no ocurre, las mejoras o cambios que se permitan en alguna que otra escuela constituirán episodios aislados son más resonancia que la que pueda lograr en su medio, y presta a quedar relegadas en el olvido cuando su promotor deje de apoyarlas.

En relación al problema que motiva este manual, no escapa al autor que la posibilidad de aplicar sistemáticamente muchas de las ideas y técnicas desarrolladas en el mismo estará condicionada a esa acción coordinada y conjunta de los miembros que integren un determinado complejo educacional. No se descuenta que las aplicaciones aisladas a cargo de docentes entusiastas de su función reditúen beneficios apreciables a su conciencia profesional. Pero una efectividad integral de lo que aquí se sugiere sólo se logrará si se dan las condiciones antes mencionadas.

Un programa de evaluación, tal como el que implícitamente se desprenderá de las páginas de este libro, alcanzará su máxima efectividad si se toman en consideración algunas o la mayoría de las siguientes proposiciones que se relacionan fundamentalmente e con los intereses y capacidades de los que se hallen interesados en montar sus funcionamiento.

a) La aplicación de nuevas técnicas o el mejoramiento de las que ya se conocen debe realizarse sobre la base de una permanente actitud científica.

Ya poco discuten la existencia de multitud de problemas en el campo educacional que sólo admiten soluciones científicas. No se duda que entre los profesionales de la enseñanza (como entre los miembros de cualquier otra actividad ocupacional) puedan existir docentes inclinados a resolver importantes problemas de su área, de modo empírico y sentimental, recurriendo al sagrado influjo de la tradición, a la opinión comprometida de los que carecen de amplitud mental y a la aceptación acrítica de esquemas desprovistos de sensatez y buen juicio, sólo porque complacen la irracionalidad de sus puntos de vista emocionales, ligados a prejuicios, ideas preconcebidas, etc.

Dentro de estos grupos se pueden incluir, por ejemplo, los que rechazan cualquier sistema de evaluación, ya que según sus expresiones, "lo que un alumno sabe y puede hacer es más importante que la nota que pueda obtener", como si ambas cosas se dieran en forma independiente.

Además ¿en qué se basan para asegurar que un alumno sabe algo y lo pueden hacer? ¿En su juicio personal? ¿Cuál es el patrón que emplean para determinar, no ya la medida, sino la propia existencia de ese saber? ¿Su propia experiencia? ¿Habrá de admitirse que puedan darse tantos patrones como experiencias existen? Naturalmente, entre los autores de temas sobre educación también se cuentan los que pertenecen a estos grupos y que a través de sus trabajos escritos pretenden resolver multitud de problemas del modo antes anotado.

Frente a estos grupos minoritarios, rémoras inconscientes del progreso, se halla una gran mayoría de auténticos profesionales del área educacional que intentan resolver sus problemas técnicos sin necesidad de caer en el uso de estereotipos o en la adopción de falsos o inconsistentes principios de autoridad, generalizaciones inadecuadas o soluciones simplistas.

La actitud científica del docente, referida a la incorporación o reajuste de técnicas y procedimientos de medición y evaluación dentro del proceso educacional, le permitirá advertir, entre otras cosas:

- *Que la evaluación es un medio y no un fin.*

En base a las conclusiones que se extraigan de un examen o a la observación directa de las conductas se tomarán o mejorarán decisiones respecto del reajuste o nuevo encauzamiento del proceso educacional ((o medidas vinculadas con otros campos: orientación, promoción, etc.).

- *Que la evaluación carece de técnicas y procedimientos infalibles*

Como en todas las áreas científicas, los errores de medición también encuentran aquí un terreno adecuado (y más fecundo por la naturaleza de lo que se pretende medir).

Si se desea comprobar, por ejemplo, la habilidad de pronunciación de un sujeto, ¿treinta palabras que incluyan las dificultades de articulación más comunes serán una muestra representativa para inferir la existencia de dicha habilidad? Equivocarse en esta apreciación ya determinará un error de muestreo. Los puntajes de las pruebas objetivas pueden estar afectados por los efectos de azar; los de las pruebas de ensayo u composición, por la subjetividad de valoración del que juzga, las observaciones directas de las conductas de los alumnos, distorsionadas por las propias antiparras mentales del observante. Saber de antemano las ventajas y limitaciones de los instrumentos y procedimientos que se deberían aplicar, facilitará la prudencia de su interpretación.

- *Que para asegurar categóricamente la efectividad de un proceso de evaluación hay que demostrarlo estadísticamente.*

Solamente una medida dirá, por ejemplo, si el examen de álgebra tomado en la víspera, posee el nivel de dificultad adecuado. La desviación Standard, si la prueba establece correctas diferencias entre los alumnos de diferente nivel de rendimiento; el coeficiente de confiabilidad, si el instrumento es consistente. (Ebel, 1965, pág.15).

b) La dirección del establecimiento y los miembros del personal que organicen un programa de evaluación eficiente, deberán coordinar sus esfuerzos.

La construcción de pruebas para comprobar resultados de la acción escolar no es una tarea sencilla ni fácil. Requieren la inversión de muchas horas de labor inteligente, no de rutina.

Si la dirección del establecimiento no ha montado una mínima infraestructura de apoyo al funcionamiento de la enseñanza, la tarea se tornará más abrumadora y menos motivable.

Parece racional pensar en que toda escuela (o grupos de escuelas) debería contar con ciertos materiales de uso común. Ya no es una utopía la posesión de un mimeógrafo (o hectógrafo), una máquina de escribir, papel y demás accesorios. Si el docente debe procurarse todos estos medios en otro lugar, la suerte de una evaluación sistemática es imprevisible.

También parece razonable liberar a los docentes de tareas menos importantes y productivas para la formación de los alumnos (algunas de las clásicas "iniciativas", episodios aislados sin conexión con objetivos claramente definidos, algunas de las largas e inútiles preparaciones de conmemoraciones, que sólo exigen sencillez y sobriedad, etc.).

La dirección de la escuela deberá crear un clima de respeto y consideración hacia el trabajo de su personal. Suele ocurrir, por ejemplo, que algún maestro confecciona un excelente fichero de problemas de aritmética para su grado. Durante todo un año, con ingenio y constancia, consigue reunir un buen repertorio que utiliza con sus propios alumnos, al año siguiente es trasladado a otro grado. El fichero ya no le servirá, y nadie lo volverá a usar. Si la dirección hubiera creado un clima de intercambio y cooperación, ese trabajo podría aprovecharlo el nuevo maestro que lo sustituya, u otros docentes que lleven idéntico grado en otras escuelas.

A su vez, él, en su nuevo grado, bien podría aprovechar creaciones efectuadas por otros colegas, este estado de cosas hace que en todas las escuelas, todos los años, estén comenzando nuevos proyectos en lugar de mantener la continuidad y mejoramiento de los ya iniciados. Calcúlese el nivel de experiencia que en muchos establecimientos se hubiera logrado si en los últimos diez años uno hubiera ensayado y confeccionado, año, tras año, nuevas técnicas para lograr una mejor ortografía, otro, nuevos métodos para resolver el problema de las diferencias individuales, o para lograr aprendizajes que perduren más tiempo, p cualquier otro proyecto de importancia real en el plano educativo.

Trasládese este punto de vista al terreno de la evaluación y se apreciará la conveniencia de aprovechar el trabajo que los docentes realicen sobre enunciados de objetivos, elaboración de ítems, análisis de resultados, etc., manteniendo los ficheros o el material confeccionado a disposición de los que quieran usarlo y aconsejando su empleo y mejoramiento, esta idea y la posibilidad de intercambiar material entre el personal interesado de escuelas primarias y secundarias, facilitará el enriquecimiento de la experiencia personal de cada uno y abreviará en mucho la dura tarea de construcción de pruebas.

c) El sistema escolar deberá disponer de algún personal más especializado en evaluación

Los capítulos que integran este manual no agotan, ni mucho menos, el tema de las modificaciones y evaluaciones escolares. Su modesto desarrollo apunta solamente a dar una idea de su ámbito y de los puntos esenciales que pueden ayudar e al maestro o al profesor. Aunque la persona que lo utilice puede perfeccionar sensiblemente sus técnicas de trabajo, siempre sería más efectivo recibir ayuda directa y sistemática (completamentada por la información bibliografía) de algún colega personal directivo o técnico de su jurisdicción. En concreto: en una escuela primaria un maestro de los grado inferiores, otro de los medios y un tercero de los superiores, deberían conocer más a fondo el problema desde el punto de vista teórico- práctico y actuar como miembros consultores de sus colegas (del mismo modo se podrían dar otras especializaciones, tales como metodología de la matemática o de los estudios sociales, psicología del aprendizaje

infantil, etc.). a su vez un grupo de tres o cuatro escuelas o un mínimo de treinta o cuarenta docentes deberán ser asesorados permanentemente por un especialista en evaluación, dependiente del organismo escolar respectivo. En la enseñanza media, salvadas las distancias de organización escolar, ocurriría otro tanto. Alguien es especial, de un grupo de profesores de matemática, por ejemplo, sería el asesor permanente de los problemas de medición y evaluación en ese campo (de igual modo en las restantes asignaturas o áreas).

Los expertos en educación (docente con amplia información práctica suficiente, acreditada por centros especializados), fijarán planes de acción, tendientes a provocar un movimiento consciente de su significación y de su repercusión en los demás aspectos del proceso educacional (objetivos, contenidos, experiencias de aprendizaje, métodos etc.).

La institucionalización de un servicio de esta naturaleza gravitaría grandemente sobre la misma empresa educacional y aceleraría su proceso de cambio hacia niveles de excelencia cada vez más elevados.

Bloque II

"LOS OBJETIVOS DE LA EVALUACIÓN EDUCATIVA"

Casonava, María Antonieta; "Ámbitos de la evaluación"
En: La evaluación educativa. España., biblioteca
Normalista. Pp. 41-47-.

CAPÍTULO II

ÁMBITOS DE LA EVALUACIÓN

Los ámbitos a los que se aplica la evaluación en el campo educativo son variados, como es obvio: programas, métodos, centros, aprendizajes, etc. Ha quedado suficientemente puesto de manifiesto en el capítulo anterior, ya que, desde sus comienzos, los modelos cuantitativos y cualitativos fueron aplicados especialmente a programas educativos que comprendían muchos elementos del quehacer docente.

No obstante –como también hemos anticipado-, la evaluación permanentemente aplicada dentro del sistema ha sido una tarea del profesor dirigida casi exclusivamente al aprendizaje de los alumnos. por tanto, los resultados del sistema educativo se han valorado de modo único en función del rendimiento que han ido alcanzando las diferentes promociones de alumnado que han pasado por él.

Evaluar un sistema educativo mediante el indicador citado –rendimiento del alumnado- es lógico y aceptable, ya que la finalidad de toda educación institucional es conseguir un buen nivel de formación en las generaciones jóvenes. Comprobar,, puntualmente y en un momento final, este resultado resulta un dato necesario para las Administraciones con competencias en la enseñanza. Lo que no parece tan razonable es que sea éste el único indicador evaluado:

- a) Porque esos resultados individuales – en la mayoría de los estudios que se llevan a cabo- dependen de la valoración realizada por diversidad de profesores, con diferentes criterios y distintos referentes para emitir los juicios evaluadores, lo que desemboca en la falta de garantía en cuanto a su equivalencia. Para que estos resultados fueran homologables sería preciso aplicar una prueba única al final de un determinado periodo de escolaridad, valorada con criterios únicos, perfectamente establecidos y, a ser posible por un equipo de evaluadores preparados especialmente para ello.

Por lo tanto, un sistema evaluado exclusivamente por los rendimientos del alumnado y en función de las calificaciones otorgadas por los profesores, no pueden ser valorados como positivo o negativo sin riesgo claro de cometer errores grandes en esa valoración.

- b) Porque los resultados que alcanzan los alumnos están condicionados fuertemente por el funcionamiento de otros muchos componentes que integran el sistema educativo y, por lo tanto, si se quiere mejorar la calidad de la enseñanza no es posible ejecutar acciones bien fundamentadas sin no se conocen las “parcelas” del sistema que funcionan adecuadamente y las que es preciso mejorar de forma clara y urgente.

El rendimiento de un estudiante no depende exclusivamente de sus capacidades o su esfuerzo personal, ni de la mayor o menor idoneidad del profesor. Depende también de la organización general de la Administración y de los centros escolares, que incidirá y repercutirá directamente en el quehacer de los planteles y departamentos de profesores y, en consecuencia, en los procesos de aprendizaje del alumnado, que mejorarán sin duda si el resto de los elementos educativos se encuentran bien regulados y ajustados a las necesidades específicas de cada situación. Como dice Fernández Pérez, M (1988, 151): « (...) sistemáticamente, durante cursos y cursos académicos, los docentes solemos “leer”, en los resultados de la evolución tradicional, los éxitos o fallos del alumno, dejando de lado todos los restantes cúmulos de variables. Las consecuencias son obvias: exactamente las mismas que tendían lugar (...) si, habiéndose pagado la luz a causa de que se han fundido los fusibles, nos empeñamos en que “lo que pasa es que está fundida la bombilla”; podemos dedicarnos a comprar lámparas nuevas, hasta centenares, y sustituirlas en nuestra escuela. La luz nunca llegará por ese camino, sólo perderemos tiempo

y dinero, a menos que nos decidamos a plantearnos el interrogante de los "porques" de los hechos observados ».

Así pues, dentro del sistema educativo aparecen tres grandes ámbitos para aplicar la evaluación:

1. La Administración educativa.
2. Los centros escolares.
3. Los procesos de enseñanza y aprendizaje.

En sentido estricto, los procesos de enseñanza y aprendizaje constituyen un indicador relevante para evaluar el funcionamiento de los centros, por lo que podría no ser consecuente reflejarlo como ámbito separado, no obstante, como son el eje en torno al cual giran todos los demás elementos y, a la vez, están condicionados por ellos y resultan sin duda, el momento clave en el que confluyen todas las normas legales, las peculiaridades personales de los alumnos, las circunstancias del entorno, las posibilidades materiales, funcionales y humanas del centro, el sistema de valores individual y social en el que se ubica, etc., creo que está justificado no sólo plantearlo separadamente, sino dedicarle una obra completa, como se hace en las páginas que siguen.

EVALUACIÓN DE LAS ADMINISTRACIONES EDUCATIVAS

El marco en el que se desarrolla el sistema educativo de un país lo determina la Administración correspondiente al emitir las normas legales que establecen las vías que guiarán su organización y funcionamiento general. Son pautas de actuación en torno a la organización del centro, sus órganos directivos y de coordinación didáctica, el modelo de síenlo curricular y los objetivos y contenidos básicos del mismo, los procedimientos metodológicos y de evaluación más adecuados para las metas que se pretenden, los criterios para la promoción y titulación del alumnado, etc. Igualmente, y en otro orden de cosas, establece los recursos económicos que se adjudicarán para la educación, a la par que los sistemas de formación inicial y permanente del profesorado, las condiciones de autorización para los centros, las titulaciones necesarias para impartir la docencia... Es decir, que en principios encontramos unas situaciones dadas que afectan al conjunto de los profesionales de la educación y que ofrecen el camino por el que debe discutir la enseñanza. A partir de esas bases, cada centro y cada profesional tomará las opciones y decisiones pertinentes y posibles para desarrollar se sistema educativo esquemático que se regula de forma legal.

La política educativa y, en consecuencia, su legislación ¿viene condicionada por elementos anteriores? Indudablemente. La estructura y principios de un sistema educativo son el reflejo de la sociedad en la que se implanta y, por lo tanto,

aparece influido por ella, por los valores que predominan y son aceptados entre sus miembros. Así, en una sociedad competitiva que valora preferentemente el éxito la posesión de bienes materiales, el sistema educativo suele proponerse dotar a los alumnos de una gran dosis de conocimientos. Conseguir personas que sean capaces de superar obstáculos ("exámenes") y que, en consecuencia, dispongan de "cabezas bien llenas". Por el contrario, si la sociedad persigue la integración personal y profesional más adecuada para sus miembros, el sistema educativo debe pretender la formación integral de la persona para que, cuando finalice su educación institucional, está dotada de un pensamiento bien estructurado, de habilidades y destrezas apropiadas para su formación continua, de actitudes participativas, creativas y críticas, de un grado de socialización adecuado, etc. En resumen, se propondrá conseguir "cabezas bien hechas" (Montaige).

No obstante, desde la educación se suele plantear una situación dialéctica y siempre problemática, pues por una parte los hechos son, a veces, tal como los acabamos de exponer, mientras que, por otra, se intenta modificar todo lo negativo en la sociedad que nos rodea proponiendo objetivos superadores que nos conduzcan a una sociedad mejor. Es decir, que el sistema educativo refleja la sociedad en la que se encuentra e intenta mejorarla paulatinamente creando en las nuevas generaciones actitudes positivas que favorezcan los avances necesarios, sin perder los bienes culturales que constituyen el patrimonio de todos.

Si realmente la Administración educativa tienen la incidencia que acabamos de comentar, se multiplica ésta cuando en un país son varias las administraciones que desarrollan las normas básicas para adaptar las a su peculiaridades, pues en cada zona territorial van a aparecer unos determinados condicionamientos y valores que indicarán en la forma expuesta anteriormente, así, el marco legal se amplía y los riesgos de sesgo hacia una tendencia de modelo social concreto, también.

Desde esta perspectiva, resulta muy importante que las administraciones que toman decisiones políticas o técnicas acerca del funcionamiento del sistema educativo en su ámbito de gestión sean conscientes de las implicaciones de su forma de actuar, es decir, que conozcan profundamente el alcance de las medidas que adoptan y sus repercusiones directas e indirectas., en síntesis, se hace imprescindible evaluar el funcionamiento general de la Administración de un país, pues podría darse el caso de que las normas y actuaciones de la misma condujeran a hechos educativos no deseados y, por lo tanto, a unos resultados deficientes del sistema en su conjunto.

Por ejemplo si la regulación es excesivamente pormenorizada produce una gran rigidez en las formas de funcionamiento de las instituciones escolares, que impiden esa adecuación curricular y de gestión necesarias para atender a la población diversa que lógicamente deberá atender; si la dirección de los centros no dispone

de un amplio margen de decisión en los órdenes que debe gestionar, no podrá llevar adelante la institución de acuerdo con sus proyectos, si faltan los recursos humanos y materiales imprescindibles para desarrollar el currículo previsto, los resultados no serán nunca los deseados, esta breve enumeración sólo es una muestra mínima de diferentes situaciones que pueden presentarse y que justifican la inclusión de la Administración como uno de los ámbitos decisivos que deben evaluarse dentro del sistema educativo.

Como este ámbito no se va a volver a tratar en la obra, pues no es el objeto de la misma, destaco a continuación algunos de los componentes que pueden o deben evaluarse (interna o externamente, según los casos) para detectar el más o menos adecuado funcionamiento de la Administración:

Estructura organizativa de la Administración

La división interna que se establece en una Administración educativa deberá estar en función de la estructura del sistema escolar, de manera que, con su funcionamiento coherente y coordinado, favorezca el adecuado rendimiento de las diversas instituciones educativas que lo componen y la máxima optimización de los recursos (de todo tipo) existentes. Por ello, para evaluar esta estructura organizativa, podrán considerarse algunos indicadores que pongan de manifiesto estos datos. A título de ejemplos, pueden citarse:

- Presenta coherencia e formal entre sus distintas unidades (Direcciones generales, Subdirecciones, servicios, Gabinetes, etc.)
- Establece la relación adecuada entre estas unidades para encauzar sus decisiones en una dirección homogénea.
- Coordina correctamente la toma de decisiones entre las diferentes unidades. Satisfacer los objetivos de cada momento.
- Mantiene los niveles de eficacia y eficiencia exigibles en función de las necesidades del sistema y de los recursos invertidos.
- Prevé las medidas y actuaciones oportunas para poseer un conocimiento permanente de la incidencia en el sistema escolar de las medidas adoptadas.
- Toma nuevas decisiones en base a los datos obtenidos de evaluación del sistema.

Normas legales emitidas

El funcionamiento del sistema educativo de un país se apoya en la normatividad que lo estructura y marca sus líneas básicas de actuación. Cualquier actividad educativa que tienen lugar está plasmada, prevista o regulada en alguna norma enmarcada sus objetivos, modo de realización, requisitos etc. La importancia de la regulación legal vienen, por lo tanto, de la incidencia que posee para favorecer o

inhibir la actividad de los centros escolares y de todas las instituciones educativas (universitaria, de investigación, etc.) según el modelo de sistema político al que responda, en función de su mayor o menor adecuación a las exigencias sociales del momento histórico en que se emiten y a la respuesta que dé a las necesidades de formación de sus nuevas generaciones.

Para llevar a acabo la evaluación de la normativa legal, se puede pensar en indicadores tales como:

- Favorecer y facilita el desarrollo adecuado del a actividad en los centros.
- Apoya el funcionamiento de la institución en sus elementos fundamentales.
- Es coherente el conjunto de normas emitidas por las diferentes unidades de la Administración.
- Es posible el cumplimiento de las mismas en su totalidad o presentan dificultades evidentes.
- Mejora la situación existente en el momento de su aparición.
- Se detectan las irregularidades más importantes que aparecen en su aplicación.
- Se corrige cuando es necesario.
- Se valora correctamente su incidencia en el desarrollo del sistema educativo.

Funcionamiento de Programas y Servicios Específicos

Habitualmente, antes de generalizar una innovación educativa, la Administración aplicada el programa correspondiente de forma experimental en determinados centros que pueden resultar idóneo para observar las ventajas e inconvenientes que presenta, corregir y reforzar los elementos necesarios y, posteriormente, generalizador al conjunto del sistema. Como es evidente, la evaluación de esta aplicación experimental es algo ineludible por necesario, ya que es el único medio del que se dispone para comprobar su proceso de funcionamiento y los resultados que se obtienen y, así, poder encauzar su futura extensión a todos los centros o situaciones equivalentes a la ensayada.

Del mismo modo, según cambian las condiciones sociales y surgen nuevas exigencias en el campo de la educación, las Administraciones van ampliando los Servicios tanto en su estructura central como periférica. Si en el primer apartado vimos la importancia de evaluar la estructura organizativa de la Administración, lógicamente las nuevas incorporaciones a la misma deben ser valoradas especialmente en sus primeros momentos de funcionamiento, pues es cuando se pueden detectar mejor las vías apropiadas para su desarrollo y consolidación posterior, a la vez que corregir los fallos que se planteen en estos primeros momentos. Esta corrección es imprescindible, pues si se afianzan determinadas

rutinas o errores desde el comienzo o de una actuación se corre el riesgo de que condicionen su evolución y futura y definitiva estructura.

Cada Programa o Servicio tendrá que evaluarse de forma prioritaria, por lo tanto, durante los primeros años de funcionamiento, con objeto de adecuar las expectativas puestas en él a su realidad, para conseguir que rinda al máximo de sus posibilidades. Algunos de los indicadores que darían cuenta de su funcionamiento, serían:

- Su funcionamiento se adecua a lo previsto y ordenado.
- Está cumpliendo los objetivos para los cuales se aplica o se pone en marcha.
- La estructura dada es la más conveniente en función de los medios con que cuenta y de lo que se debe conseguir.
- Los elementos curriculares del programa resultan coherentes en su aplicación.
- Están apareciendo dificultades significativas¹ en su implantación, desarrollo o funcionamiento general.
- Se detectan valores especialmente positivos² durante su aplicación experimental por la incidencia específica de su funcionamiento.
- Se están tomando las medidas correctas adecuadas para mejorarlo.
- Se prevé que no será posible su extensión o que es preciso suspender su funcionamiento.

EVALUACIÓN DE LOS CENTROS ESCOLARES

Los hechos educativos tienen lugar en los centros escolares o instituciones educativas en general., más concretamente, los procesos de enseñanza y aprendizaje se desarrollan fundamentalmente dentro del aula. Son la consecuencia de la interacción permanente entre profesor/alumno y alumnos/alumnos, de la comunicación formativa que se produce en toda situación educativa por excelencia.

En los centros escolares concluyen todo los agentes externos antes citados, con los puramente técnicos e internos propios de la educación institucional. Es decir, en ellos hay que aplicar las normas que enmarcan el funcionamiento general del sistema educativo (procedentes de una línea política determinada), a atenerse estructural y funcionalmente a la legalidad vigente, atender a los diversos grupos socioculturales que pueden coincidir e integrarse en cada momento y desarrollar los procesos de enseñanza y aprendizaje más adecuados a la situación, teniendo en cuenta las premisas y condicionamientos anteriores.

¹ Será necesario detallarlas para evaluar cada una de ellas.

² Será necesario detallar para poder evaluar su incidencia.

Es una situación realmente compleja la que se plantea a los educadores y directivos de los centros para hacer funcionar este conglomerado de circunstancias obteniendo el rendimiento que la sociedad espera de ellos en cuanto a la formación e integración personal y laboral de las generaciones jóvenes.

Parece, por ello, que el control y evaluación de su funcionamiento deberían constituir una práctica habitual dentro de todo sistema educativo. El comprobar y valorar cómo van engranando los diferentes y muy diversos componentes que, integran es decisivo para su buena marcha, sobre todo, para procurar su mejora paulatina tanto en sus elementos de administración como en los más estrictamente pedagógicos. Ésta será la única forma de alcanzar unos resultados congruentes con las expectativas que la sociedad pone actualmente en los sistemas educativos institucionales. La evaluación se conforma, así, como un elemento primordial de perfeccionamiento de los centros escolares, al incorporarla a los procesos de todo tipo que se producen en ellos y ofrecer los datos suficientes y fiables para aplicar permanentemente las medidas que modifiquen o refuercen lo necesario.

No obstante esta evidencia palpable, hay que decir que en España no está generalizada la práctica evaluadora en los centros educativos. No es que la Administración haya dejado de ejercer su función de control a través de la supervisión educativa –al igual que ocurre en cualquier país y con cualquier modelo de gobierno–, sino que hasta estos últimos años no se han comenzado a aplicar un sistema de evaluación técnica en consonancia con las virtualidades que ésta posee y con las evidencias de mejora que se plantean desde los sectores sociales.

Tampoco ha sido, hasta ahora, una práctica habitual en los centros su evaluación interna, con la rigurosidad y sistematicidad que resulta imprescindible, a pesar de que sus directivos y profesorado debieran ser los primeros interesados en ofrecer una educación de calidad al alumnado que atiende y, por lo tanto, en valorar hasta qué punto lo están consiguiendo y cómo mejorar día a día su labor y el aprendizaje de los niños y jóvenes que se están formando.

Como este apartado de la obra lo único que pretende es enmarcar la evaluación de la enseñanza y el aprendizaje dentro de los ámbitos que debe cubrir la evaluación dentro de otro sistema educativo, no entro a desarrollar la evaluación de centros. Sí quiero insistir en que el indicador principal de su buen funcionamiento y, por lo tanto, el referente que ningún equipo directivo ni profesor debe perder al evaluarlo, es el rendimiento que alcanza su alumnado³. Por ello, la evaluación de los procesos educativos podría incluirse dentro de la evaluación del centro, como

³ Entiéndase rendimiento en su sentido más amplio desde el punto de vista educativo, no sólo como éxito en las calificaciones obtenidas, es decir; rendimiento como logro de una formación integral, completa..., que en unos casos se traducirá en la posesión de un título como expresión social de la formación alcanzada, y otras no, pues son diversa las vías de la persona para llegar a su mejor modo de vivir en sociedad con una integración positiva para todos.

uno de sus componentes. Si la aislamos es por poseer especificidad suficiente como para ser objeto de la obra completa.

Por lo que se refiere a la evaluación externa de los centros, cabe destacar – en España- la aplicación que se ha realizado, de forma sistemática, del *Plan de Evolución de Centros* por parte del Ministerio de Educación y Ciencias en su ámbito de gestión. Se ha dirigido a todo tipo de centros y se ha aplicado en 1.07 a lo largo de cinco años (V.V.AA.: 1996). su finalidad ha sido eminentemente formativa- a pesar de ser evaluación externa- y se ha contado, para su realización, con la colaboración de toda la comunidad educativa (figura 1). Otro modelo de evaluación externa de centros lo Sustituye el de auditoria, propuesto por Martínez Aragón, L. y Pérez Juste, R. (1989, 177-263), en el cual se lleva a cabo un análisis exhaustivo de sus funcionamiento mediante la consideración de todos sus componentes a través de un número amplio de indicadores y controles internos que den validez al modelo.

En cuanto a la evaluación interna de los centros, me parece especialmente importante porque es el medio más eficaz de mejorar, paso a paso, su calidad de funcionamiento y resultados. Los protagonistas de la acción educativa son, los primeros que deben evaluarla; recopilar datos de modo riguroso y sistemático, analizados, obtener unas conclusiones, valorarlas y tomar las decisiones necesarias para perfeccionar su proceso y su rendimiento. Si esto se lleva a cabo, las medidas de mejora son inmediatas, no es preciso esperar a las conclusiones de una evaluación externa que, además, sólo es posible cada cierto número de años, nunca va a ser continuada, como resulta imprescindible en todo desarrollo procesual humano siempre perfectible. Por otra parte, la complementariedad de las informaciones obtenidas externa e internamente es otra garantía de fiabilidad para llegar a conclusiones correctas, pues se enriquecen mutuamente y facilitan la discusión, el debate y la mejor solución para las cuestiones planteadas.

Figura 1
Modelo de ficha del Plan de evaluación de Centros (M.E.C.)

Ficha de evaluación no. 14 (Equipos docentes II): b) Departamentos de Orientación		4.3 Equipos docentes
Denominación del Centro:		Clave:
Localidad	Nivel: Educación Secundaria	Régimen jurídico:
1. Número promedio de reuniones cada mes.	012345	Todos los componentes del departamento de orientación participan activamente en su funcionamiento.
2. El grado de participación de los componentes en el funcionamiento de Orientación es muy bajo.	012345	El jefe del departamento de orientación actúa en estrecha colaboración con el Equipo directivo.
3. El jefe del Departamento de Orientación actúa individualmente y sin relación alguna con el equipo directivo.	012345	La actividad del departamento de orientación se ajusta planamente a lo previsto en el plan de actividades y sus revisiones razonadas.
4. La actividad del departamento de Orientación no se ajusta a lo previsto en el Plan de actividades.	012345	El departamento de orientación coordina

5. El Departamento de orientación no coordina la orientación educativa.	012345	adecuadamente la orientación educativa especialmente en los cambios de ciclo o etapa.
6. El departamento de orientación no coordina la orientación educativa para la elección entre las opciones académicas formativas y profesionales.	012345	El departamento de orientación coordina adecuadamente la orientación educativa para la elección entre las opciones académicas y profesionales.
7. El departamento de orientación no realiza ningún tipo de evaluación psicológica de los alumnos que lo necesitan.	012345	El departamento de orientación realiza eficazmente la evaluación psicológica y pedagógica de los alumnos que lo necesitan colaborando con los profesores en la prevención y de lección de problemas de aprendizaje.
8. El departamento de orientación y los tutores no desarrollan el plan de orientación y el de acción tutorial previstos en el P. C. E.	012345	El departamento de orientación y los tutores desarrollan de forma coordinada y eficaz el plan de orientación y el de acción tutorial previstos en el P. C. E.
9. El departamento de orientación no realiza actividades educativas y adaptaciones curriculares para los alumnos con problemas de aprendizaje.	012345	El departamento de orientación realiza actividades educativas y adaptaciones curriculares adecuadas para alumnos con problemas de aprendizaje.
10. El departamento de orientación no da respuesta educativa a los alumnos que necesitan seguir programas de diversificación curricular o de garantía social.	012345	El departamento de orientación da respuesta educativa eficaz a los alumnos que necesitan seguir programas de diversificación curricular o bien programas especiales de garantía social.
11. El departamento de orientación no interviene en la elaboración del consejo orientador sobre el futuro académico y profesional del alumno que no alcanza los objetivos de la etapa.	012345	El departamento de orientación participa activamente en la elaboración del consejo orientador sobre el futuro académico y profesional del alumno que no alcanza los objetivos de la etapa.
12. El departamento de orientación no coordina la orientación laboral y profesional con otras Administraciones.	012345	El departamento de orientación coordina eficazmente la orientación laboral y profesional con otras Administraciones o Instituciones.
13. Los alumnos del centro desconocen el programa del departamento de orientación.	012345	

En este sentido aparecen publicadas⁴ varias propuestas para la evaluación interna de los centros docentes:

Pere Darder y José Antonio López (1985) ofrecen en el Quafe-80⁵ un cuestionario para la reflexión de los profesionales que trabajan en el centro, en torno a 26 amplios componentes. Se describe el componente como encabezamiento de la reflexión y, a continuación, se escalonan cinco situaciones- de peor a mejor- para que cada profesional señale la que resulta más similar a la de su centro, según su propia visión. En la misma publicación se dan indicaciones para la aplicación y obtención de conclusiones de este modelo evaluador. La adaptación de este cuestionamiento a los centros de Educación Infantil (ECEI), coordinada por Pere Darder y Joan Mestres (1994), sigue la misma pauta de trabajo e incorpora un cuestionamiento dirigido a las familias y unos registros para la evaluación cuantitativa del centro (figura 2) .

⁴ Selecciono las de autores españoles que considero con posibilidad de aplicar en los centros y fáciles de encontrar por ser todas relativamente recientes.

⁵ Dada la fecha de publicación de la obra. La terminología que se utiliza en ella no se adecua a la que ahora se maneja, en algunas de las situaciones que se plantean. No obstante, continúa siendo de interés, aunque precisaría una actualización para adaptarla, fundamentalmente, a la nueva nomenclatura.

Miguel Ángel Santos Guerra (1990) presenta un modelo de evolución etnográfica, cualitativa, eminentemente procesal y cuya forma de aplicación concreta debe decirse cada equipo de docentes en función de las características de su centro. Toda organización educativa supone algo vivo, y cambiante, y, por ello, aquí se ofrecen comentarios teóricos además de las descripciones de técnicas e instrumentos que pueden ser útiles para definir el modelo de evaluación más adecuado en cada momento o situación.

Una propuesta personal queda recogida en *La evolución, garantía de calidad para el centro educativo* (1992), obra en la cual se hacen planteamientos teórico-prácticos abiertos, creo que con una aplicabilidad asequible, de manera que sea posible que los equipos directivos y docentes puedan organizar su evaluación interna en función de sus necesidades y capacidad de acción. Las técnicas e instrumentos de evaluación se describen con el suficiente detalle como para que sea factible su aplicación o elaboración (figura 3). Ello se facilita con la enumeración de los componentes evaluables en un centro y con la formulación de indicadores- a título de ejemplo- para cada uno de ellos. Para finalizar, se presentan tres posibles modelos de estructurar la evaluación del centro: evaluación sectorial del centro (de uno o varios componentes del mismo), evaluación global del centro mediante los mecanismos previstos en la normativa legal vigente y evaluación de la implantación de un programa, en todos los casos, la propuesta es cualitativa, etnográfica y, por lo tanto, flexible, para que pueda ser utilizada, como antes señalo, de acuerdo con las finalidades u opciones que el centro se marque . Otra alternativa, más reciente, es la que propongo intentando enlazar la evaluación que se realice con el plan de mejora inmediato para el centro (Casanova, M. A.: 1997). Sintetizo los aspectos más aclaratorios de la misma:

FIGURA 2

Modelo del planteamiento presentado en el ECEI, referido al componente «Coordinación»

19. Coordinación

Cuando un conjunto de personas o una institución quieren realizar un proyecto común resulta necesario coordinar orientaciones, procedimientos de trabajo y actitudes concretas. La falta de coordinación pondría en cuestión la posibilidad de realizar una tarea en común e y de dudar de la intención de realizarla. Coordinar quiere decir velar para que en la práctica se trabaje en común, en función de un proyecto.

Es importante revisar si se da o no la coordinación, a quien se confía esta tarea, de qué forma se implican las diversas personas, qué funciones se les encomienda, etc.

- A. Por diferentes razones, no se llega a un trabajo y una orientación coordinada entre los maestros de la escuela. Puede haber coherencia personal en la actuación de cada uno, pero la coordinación entre las tareas de los diferentes maestros no es una preocupación que lleva a realizaciones concretas.
- B. Los esfuerzos de algunos maestros o de los órganos de gestión consiguen que sectores o grupos de la escuela actúen coordinadamente en hechos puntuales, campos concretos o durante un tiempo determinado. Este esfuerzo no tienen continuidad por diversas causas (cambios de personas, disponibilidad de tiempo, poca institucionalización, etc.).
- C. En la escuela hay preocupación generalizada y un esfuerzo por conseguir una acción coordinada, a partir de unos acuerdos concretos en el proyecto de escuela. A pesar de esto, se detectan zonas importantes de dispersión en la actuación que hacen pensar en la necesidad de reestructurar las formas de coordinación.
- D. En la escuela están bien definidos los canales formales que permiten realizar coordinadamente los diversos aspectos del proyecto. Esto hace que, en la práctica, los esfuerzos y las aportaciones individuales tengan una repercusión positiva en el funcionamiento general de la escuela. Se controla la coordinación existente y la eficacia de los canales de coordinación.
- E. La necesidad de encontrar un equilibrio entre la existencia de una orientación común en el centro y el reconocimiento de unos márgenes de actuación personal, orienta la revisión de los ámbitos y mecanismos de coordinación. Se introduce innovaciones a partir de los resultados de la revisión de la profundización en el análisis de la institución.

FIGURA 3

Modelo de escala ofrecido en la evaluación, garantía de calidad para el centro educativo

(Escala para valorar el nivel de expectativas del profesorado respecto a los alumnos y alumnas)

INDICADORES	Siempre	En bastantes ocasiones	Algunas veces	Nunca
1. Sus alumnos, mayoritariamente, son capaces de alcanzar los objetivos previstos en el PC.				
2. Sus alumnos participan activamente en las actividades escolares que se realizan.				
3. Sus alumnos participan activamente en las actividades extraescolares que se realizan con ánimo de superación.				
4. El PC recoge actividades complementarias (cognitivas, afectivas, motrices) para que los alumnos puedan alcanzar el máximo desarrollo posible.				
5. Usted pone de manifiesto ante sus alumnos la colaboración que espera de ellos para que se cumplan las normas de convivencia del centro.				
6. Usted expresa a sus alumnos la confianza que tienen en que superen los objetivos del curso.				
7. en sus clases se trabaja con interés en las actividades que se proponen.				
8. sus exigencias son elevadas, aunque ajustadas a las capacidades y, los intereses individuales de sus alumnos.				
9. Usted, como profesor, valora más el esfuerzo realizado por sus alumnos en función de sus capacidades, que el éxito final.				
10. Al final de cada actividad, usted da oportunidad a sus alumnos a de autoevaluar sus trabajo.				
11. Al final de cada actividad, usted da oportunidad a sus alumnos de coevaluar el trabajo desarrollado por todo el grupo.				
12. Usted, en su intervención como profesor en el aula, facilita la crítica constructiva y superadora por parte de sus alumnos.				
13. en el centro se organizan actividades culturales complementarias, con intervención activa de todos los sectores de la comunidad escolar.				
14. Sus alumnos valoran positivamente los aprendizajes que adquieren en el centro.				

15. sus alumnos valoran los aprendizajes porque comprenden la repercusión que tendrán en su vida futura.				
16. Sus alumnos asisten al centro con gusto y aceptación del mismo.				
17. Los alumnos están representados en el Consejo Escolar del centro.				
18. Los alumnos intervienen oportunamente u correctamente en la resolución de las cuestiones que se plantean en el Consejo escolar.				
19. En el conjunto del profesorado existe la opinión de que el centro funciona satisfactoriamente.				
20. La tasa de superación de los cursos, por parte del alumnado, es clara.				
21. Los alumnos, cuando pasan a un nivel educativo superior a o a un puesto de trabajo, son capaces de rendir adecuadamente.				

En las teorías sobre «modelos ideales», se pretende llegar a definir indicadores de calidad que marquen definitivamente cómo debiera funcionar un sistema educativo con garantías de respuesta adecuada a las demandas de formación que en cada momento exige la sociedad. Partiendo de ese modelo, debe calcularse matemáticamente la distancia existente entre los ideales propuestos y el funcionamiento real del centro. Cuanto más ajuste se aprecie entre ambos, mayor será la efectividad educativa y viceversa.

Para poder aplicar estas opciones – en principio de carácter matemático- a la realidad educativa, hay que admitir que es necesario aprender a trabajar en un ámbito de razonable relatividad, como ocurre en todas las ciencias humanas, y en particular, en la educación. Por suerte, las personas no somos exactamente cuantificables ni medibles y, por lo tanto, habrá que trabajar con la mayor exactitud posible, con el mayor rigor del que seamos capaces, pero sin pretender constreñir el desarrollo educativo/ pedagógico a modelos exclusivamente matemático/estadísticos.

En ese sentido, tomo la idea básica anteriormente citada y la desarrollo, a continuación, en términos de cualidad y descripción. Es posible llegar a un modelo evaluador (diagnóstico, formativo, verificador) individualizando para cada centro, elaborado por los profesionales que trabajan en él y sumamente práctico para dirigir el cambio hacia la mejora en todos los órdenes de funcionamiento.

Con objeto de llevarlo a cabo, es preciso desarrollar las siguientes fases:

- a) Decidir los componentes del centro que se van a evaluar, determinando las prioridades y temporalizando su realización.
- b) Para cada componente:
 - ❖ Definir el objetivo general que se pretende alcanzar.
 - ❖ Formular indicadores (1) que sirvan de pauta cercana, bien definida, evaluable para valorar procesualmente el camino hacia ese objetivo y llegar a concluir, con rigor, los aspectos conseguidos y los puntos débiles del mismo.
 - ❖ Examinar la realidad del centro= individual o colectivamente en relación con cada indicador ®
 - ❖ Determinar las distancia entre el indicador y la realidad (Distancia I/R).
 - ❖ Deducir las acciones necesarias para llegar a cubrir las distancias identificadas (Actuaciones para el plan de mejora).

En la figura 4 se representa un desarrollo parcial de esta propuesta, como ejemplo que facilita la visualización de lo expuesto. Si se llega a disponer de una buena síntesis de “distancias” existentes en el centro para llegar a sus metas, se dispondrá ya del camino –del plan de mejora- que debe recorrerse para alcanzarlas. Y ese camino constituye el quehacer diario del centro, por lo cual conviene marcar específicas necesarias que habrá que realizar y que faciliten su aplicación, seguimiento y valoración consecuente.

Quiero dejar reseñada, igualmente, la publicación del Ministerio de Educación y Ciencia sobre evaluación interna del centro educativo (Casanova, M.A. y otros: 1993) incluida en el material elaborado para el curso dirigido a formación de equipos directivos. Creo que también resulta útil, pues ofrece, al igual que en las anteriores, diversas posibilidades de aplicación,, con modelos de instrumentos concretos y, sobre todo, en este caso ¿, una gran cantidad de indicadores de calidad para llevar a cabo la evaluación inicial, procesual y final de los diferentes componentes del centro cuya evaluación se propone en la obra.

En el momento en que aparece la segunda edición de esta obra, el Ministerio de Educación y Cultura está proponiendo la aplicación del modelo de la Fundación Europea para la Gestión de Calidad (EFQM), que, procedente del mundo empresarial, parece mostrarse válido en el mundo de la educación. Se considera un instrumento de máxima utilidad para mejorar el funcionamiento de los centros, y centra su atención en nueve criterios, estimados como las claves para la efectividad y mejora permanente de la institución que lo aplica: liderazgo, estrategia y planificación, gestión de personal, recursos, procesos, satisfacción del cliente, satisfacción del personal, impacto en la sociedad y resultados del centro.

La evaluación continuada de estos criterios, mediante subcriterios señalados al efecto para cada uno de ellos, permite detectar los puntos fuertes y las áreas de mejora, por lo que constituye la base del perfeccionamiento que pueda producirse con su desarrollo y a aplicación (MEC, 1997).

Hay varias obras más publicadas en estos momentos, que reseño en a bibliografía. He comentado las que considero más adecuadas –por la filosofía de la que parten y las posibilidades de aplicación que ofrecen- a la situación real y actual de los centros, que deben comenzar la práctica evaluadora partiendo casi de una nula experiencia en este sentido. Constituye yu conjunto de métodos que pueden favorecer la reflexión por parte de los profesionales- y la comunidad educativa en su conjunto, si así se plante- ayudar a que se implante sistemática y rigurosamente para llegar a su utilización en orden a la mejora continua de los procesos educativos que tienen lugar en los centros y a su buen funcionamiento general, así como, consecuentemente, aun óptimo rendimiento en la formación del alumnado.

Para finalizar, dentro del planteamiento abierto que estoy presentando para la evaluación del centro, quiero dejar constancia, igualmente, de que, con las pautas que se ofrecen en bibliografía citada en cuanto a fases de la evaluación, participantes, elaboración de instrumentos o definición de técnicas , el profesorado de un centro puede organizar su propio modelo de evaluación definiendo:

- a) Para qué evaluar (objetivos de la evaluación)
 - b) Qué se va a evaluar (objeto de la evaluación)
 - c) Quiénes participarán y de que manera
 - d) Fases de la evaluación y temporalización de las mismas.
 - e) Metodología de la evaluación.
- Técnicas que se utilizan para la evaluación
 - Elaboración o selección de los instrumentos adecuados para ala recogida de datos e información final.
- f) Obtención de conclusiones finales, toma de decisiones e información a los afectados.

En relación co0n el posible modelo de instrumentos en los que se plasmen los objetivos evaluables y los indicadores adecuados para ello, puede procederse siguiendo la pauta que se presenta en la figura 4ª. Se secciona el objetivo que se pretende conseguir y que se evaluará y se describe el grafo máximo de idoneidad que parece oportuno para el centro. A partir de ahí, se formulan los indicadores operativos, con posibilidad de valorar concretamente mediante las técnicas previstas, y se organizan las acciones que es preciso llevar a cabo para favorecer la captura de información de modo riguroso y sistemático. Según se van

comprobando procesualmente los datos necesarios, se cumplimenta la escala de la figura 4ª y se analizan e interpretan los mismos de acuerdo el referente de idoneidad descrito anteriormente.

FIGURA 4

De la evaluación al plan de mejora

Componente evaluado: Departamentos didácticos/ Equipos de ciclo
Objetivos: el profesor participa activamente en el funcionamiento del departamento/ equipo / equipo de ciclo. Se colaboran en la elaboración y aplicación de las programaciones de área/aula. Se llevan a cabo actuaciones encaminadas al perfeccionamiento científico. Se actúa con criterios homogéneos en relación con los alumnos y sus familias.

Indicador	Realidad del centro	Distancia I/R	Actuaciones para el Plan de Mejora
1. El equipo/departamento se reúne periódicamente.	Está establecido un calendario de reuniones, pero no se cumple estrictamente. En el trimestre se programaron 6 reuniones y se celebraron 4.	-Falta la celebración de 2 reuniones.	1. Llevar a cabo todas las reuniones que se programen.
2. Asiste todo el profesorado a las reuniones	Asiste la mayoría, pero sólo se han celebrado dos reuniones con todos los miembros del equipo departamento.	- han faltado 6 profesores a las reuniones del trimestre.	1. Plantear la situación en la primera reunión del próximo trimestre.} 2. Llegar a conclusiones acerca de la necesidad de la asistencia de todos, para que sean efectivos los acuerdos y el trabajo. 3. Asistir todos a todas las reuniones.

Indicador	Realidad del centro	Distancia I/R	Actuaciones para el Plan de Mejora
3. Se tomen acuerdos comunes	Se adoptan decisiones, pero casi siempre decide una minoría porque muchos profesores se inhiben, no se comprometen en ellas.	- Falta la implicación real del 70% del profesorado	<ol style="list-style-type: none"> 1. Plantear la situación en la primera reunión del próximo trimestre. 2. Analizar las causas de la situación. 3. Concretar claramente los problemas existentes. 4. Proponer un plan progresivo para ir superando las dificultades de cada profesor. 5. Llegar a tomar acuerdos con el compromiso de todo el profesorado.
4. Se reflejan los acuerdos en actas documentos equivalentes.	Están regidos los acuerdos adoptados en cada reunión en un libro de actas.	- Ninguna	<ol style="list-style-type: none"> 1. Continuar como hasta ahora.
5. Se elabora en común la programación	Existe, por escrito, una programación del área/s, elaborada en su momento por un pequeño equipo y revisada parcialmente cada año, no está muy claro que todo el profesorado del departamento/equipo la respeta con sus respectivos grupos de alumnos.	<p>-Falta la revisión y actualización de la programación existente por parte de todos los miembros del departamento equipo.</p> <p>-Falta la implicación de los miembros del departamento/equipo en la elaboración de la programación.</p> <p>-Falta la seguridad de que el profesorado respeta los acuerdos tomados con cada uno de los grupos de alumnos/as.</p>	<ol style="list-style-type: none"> 1. Leer individualmente la programación existente. 2. Comparar entre las programaciones reales del profesorado y la formalmente adoptada. 3. Ajustar en común la programación. Acuerdo para llevarla a cabo, con las adaptaciones que requieran determinados grupos, en su caso. 4. Revisarla trimestralmente y evaluar su aplicación. Modificar los elementos necesarios, para su mejor

			adecuación al proyecto curricular y al alumnado.
6. Se tienen de criterios de evaluación comunes.	En la programación existen criterios de evaluación que debe aplicar todo el profesorado. No obstante, no existe constancia clara de que esto sea así.	-Está pendiente la discusión, ajuste y aceptación de los criterios por parte de todo el profesorado. - La definición, poco concreta, de los criterios, dificulta la homogeneidad en su aplicación. -Como consecuencia, no hay seguridad de que se esté evaluando al conjunto del alumnado con la misma pauta.	1. Revisar y discutir los criterios de evaluación del área, una vez acordada la programación. 2. Formular los criterios con términos evaluables, concretos, entendidos de la misma forma por todos y asumidos, también, por el departamento/equipo. 3. Secuenciar los criterios de evaluación por ciclos o cursos, en su caso, y por trimestres, para facilitar su aplicación homogénea. 4. Revisar trimestralmente la adecuación de los criterios establecidos para que los informes de evaluación resulten equivalentes. 5. Ir ajustando los criterios progresivamente.
7. La información a las familias se realiza con iguales criterios.	Los informantes que se dan trimestralmente a las familias responden a un modelo común y a los criterios establecidos en el centro para realizarlos.	-Ninguna. -No obstante, cuando se revise y ajuste el indicador anterior, si se detectan la necesitada de modificar los criterios de evaluación, habría que revisar los informes que se facilitan a las familias.	1. Se determinarán las actuaciones necesarias en función de los resultados del indicador anterior.

Indicador	Realidad del centro	Distancia I/R	Actuaciones para el Plan de Mejora
8. El profesorado se reúne con los padres y las madres de los alumnos.	Cada profesor tiene un día semanal señalado para recibir a las familias., y se cumple con regularidad.	-Ninguna, en relación con lo establecido. Se está detectando la necesidad de mantener una reunión con los padres al comienzo del curso, para darles información acerca de las decisiones adoptadas para el funcionamiento general del curso.	1. Parece interesante añadir la reunión con los padres al comienzo del curso, por lo que se planificará para el próximo.
9. Se seleccionar el material didáctico para uso del alumnado.	Cada profesor selecciona el material, pero no se hace en común, por parte del conjunto de profesorado del departamento/equipo.	-Falta el trabajo en común del profesorado para decidir cuál es el material didáctico más idóneo para el alumnado y establecer criterios para su utilización	1. Dedicar una reunión a planificar la organización del trabajo pendiente. 2. Examinar el material existente en el centro y revisar catálogos u ofertas de nuevo material interesante para el área. 3. Seleccionar, en dos sesiones como máximo, el material adecuado para utilizar en el centro. 4. Establecer criterios de organización, en caso de que sea necesario compartir el material.
10. se tienen previsto el horario para atender las consultad del alumnado.	Semanalmente, cada profesor tiene horas previstas para atender al alumnado fuera del aula.	-Ninguna	1. Continuar como hasta ahora.

Indicador	Realidad del centro	Distancia I/R	Actuaciones para el Plan de Mejora
11. Se mantienen relaciones con otros departamentos didácticos o equipos de ciclo.	Hay reuniones quincenales interciclos, para dar cohesión a la programación vertical del área en el centro. Entre los departamentos didácticos es nula la relación.	-Falta la coordinación entre el final de Educación Infantil y el primer ciclo de Educación Primaria.	<ol style="list-style-type: none"> 1. Establecer un calendario de reuniones entre el/ la coordinador/a de Educación Infantil y el/ la del primer ciclo de Educación Primaria. 2. Establecer un calendario de reuniones para comenzar a trabajar entre el tercer ciclo de Educación Primaria y el primero de Educación Secundaria. 3. Organizar una reunión con todos los jefes de departamento de Educación Secundaria, para tratar acerca de la conveniencia de establecer relaciones con objeto de conexionar los contenidos comunes y programar las áreas atendiendo la interdisciplinariedad. 4. Hacer un seguimiento detallado de estas primeras reuniones, determina sus conveniencia y regulariza su funcionamiento y su incidencia en las diferentes programaciones.

Indicador	Realidad del centro	Distancia I/R	Actuaciones para el Plan de Mejora
12. Se realiza actividades interdisciplinarias, con la colaboración de otros departamentos.	No se realiza ninguna actividad de este tipo, como queda reflejado en el indicador anterior.		-Falta establecer las reuniones de trabajo necesarias para poder aplicar una programación interdisciplinar.
13. Se cuenta con materias para la formación permanente de sus con ponentes	Existe una biblioteca ene. Centro, pero no se actualiza convenientemente. Quizá por eso se utiliza poco por parte del profesorado. No existe otro tipo de materiales.	-Actualizar la biblioteca. -Incorporar otros materiales ahora existentes para diferentes áreas (informático, musical, plástico, etc.).	1. Revisar catálogos de interés APRA la actualización científica y didáctica del profesorado., 2. Comparar los libros que resulten más interesantes y urgentes. 3. Seleccionar otro tipo de material. 4. Establecer prioridades para su adquisición progresiva.
14. El conjunto del profesorado asiste a actividades de formación permanente.	Algún profesor ha asistido en el último año a cursos de actualización, pero no es una situación general no constante.	-Falta la detección de necesidades de formación del profesorado del ciclo/departamento. -Asistir a los cursos a o actividades que respondan a las necesidades de formación.	1. realizar la detección de necesidades de formación en los diferentes departamentos didácticos, median una encuesta y una propuesta en común. 2. Seleccionar, de entre la oferta existente, los cursos u otros tipos de actividades que resulten interesantes, en base a los datos obtenidos en el punto anterior.

Indicador	Realidad del centro	Distancia I/R	Actuaciones para el Plan de Mejora
			3. Elaborar un calendario par la asistencia priorizada del profesorado a las actividades de actualización previstas. 4. En su caso, organizar un proyecto de formación en el propio centro, que cubra factores generales de interés común.
15. Se lleva a cabo algún programa experimental par ala mejora de la enseñanza en el área correspondiente.	En el ámbito de un departamento y un ciclo se lleva a cabo una experiencia de innovación didáctica.	-Falta extender experiencias a otros departamentos y ciclos, de manera que su incidencia sea lago mayor en el funcionamiento general del centro.	1. Examinar en qué +áreas puede ser más urgentemente o interesante introducir innovaciones. 2. Preparar un discerní de dos experiencias más para el curso próximo.
16. Se organizan actividades extraescolares dirigidas a la comunicación educativa.	Las actividades extraescolares las organiza un grupo de profesores en colaboración con el consejo escolar.	-No está claro que estas actividades deban partir de los departamentos. Si acaso, sería interesante su propuesta.	1. En principio, d se decide continuar como hasta ahora, admitiendo propuestas de los departamentos y ciclos.
(Este ejemplo está realizado a partir de la figura 4 de Casanova , M.A. (1992, 96), donde aparece una lista de control para evaluar los departamentos didácticos).			

Esta maneta de concretar un plan de evaluación anual para el centro, permite acomodarla a las posibilidades de cada momento, en función de la formación de los integrantes del plantel y de otras circunstancias.

FIGURA 4ª*Modelo de registro para evaluar los componentes del Centro.*

Currículo. Proyecto curricular definido y elaborado por el conjunto del profesorado.

Nivel máximo de idoneidad. El profesorado adapta y elabora el currículo escolar para los diferentes ciclos o cursos de la etapa educativa.

Indicadores de calidad	Si/Siempre	Mucho/ En muchas ocasiones	Poco/ En pocas ocasiones	No nunca
<ol style="list-style-type: none"> 1. El currículo para cada ciclo o curso está adaptado a las características ambientales del centro y a las peculiaridades de su alumnado. 2. El currículo de cada ciclo ha sido desarrollado por el equipo docente de este ciclo. 3. el currículo del ciclo fue elaborado participativamente con el equipo directivo, la comisión pedagógica, los padres y los alumnos. 4. cada profesor adapta a su grupo de alumnos el currículo del ciclo. 5. El currículo abarca todas las áreas de aprendizaje del ciclo. 6. El currículo integra objetivos para el desarrollo intelectual, psicomotor, afectivo, social, de actitudes, hábitos y técnicas de estudio, interés, etc. 7. Los objetivos de aprendizaje establecidos resultan significativos para el alumno. 8. Los objetivos de aprendizaje establecidos son adecuados para el nivel de desarrollo de, los diferentes grupos de alumnos. 9. los aprendizajes propuestos para los ciclos que integran la etapa son coherentes en su desarrollo vertical. 10. todo el profesorado trabaja en equipo para lograr la continuidad en el aprendizaje del alumnado. 				

EVALUACIÓN DE LOS PROCESOS DE ENSEÑANZA Y APRENDIZAJE

Las páginas que siguen se centran en la evaluación de los aprendizajes de los alumnos y alumnas, así como de los planteamientos de enseñanza que realiza el profesorado. Partirá de los principios básicos que justifican la propuesta que se ofrece, y de la necesidad de modificar radicalmente el modelo tradicional de

evaluación, fundamentando esencialmente en la aplicación de pruebas puntuales que pueden valorar lo memorizado hasta ese momento, pero que en ningún caso *garantizan* que el aprendizaje está realizado correctamente y, por tanto, incorporado a las estructuras mentales del alumno y actuando como generador de nuevos aprendizajes y estructuras.

La propuesta de evaluación será abierta, cualitativamente y ampliadora del modelo actual. Espero que resulte superadora de algunos prejuicios del modelo actual. Espero que resulte superadora de algunos prejuicios y condicionamientos que aparecen actualmente en el campo de la evaluación y que contribuya, de este modo, a la aplicación de un modelo formativo y continuo integrado en el desarrollo del currículo, colaborando a la mejora del mismo y de los propios procesos de aprendizaje del alumnado.

SANTOS, Guerra Miguel A. “Patología general de la evaluación educativa” En: la evaluación: un proceso de dialogo, comprensión y mejora. Granda Ed. Aljibe pp 15-31.

PATOLOGÍA GENERAL DE LA EVALUACIÓN EDUCATIVA *

El desarrollo curricular, en cualquiera de sus niveles, está recorrido por un proceso evaluar de triple naturaleza: la evaluación diagnóstica, la evaluación procesual y la evaluación de término. En el buen entendido de que el proceso es circular, meramente lineal, dinámico en su entraña.

La toma de decisiones, inicial o de proceso, nace de una valoración precisa y de un análisis del planteamiento, de la acción, del contexto y condiciones de la misma y –en su caso- de los resultados.

En el complejo mundo de la educación, por tantas motivos particular u cambiante, se realizan diversos procesos evaluadores del currículo que están afectados de las patologías más diversas.

Cada uno de los desórdenes que afectan a este proceso (en el marco de referencia macrocurricular = sistema, mesocurricular= centro y microcurricular= aula) puede ser estudiado en sus signos, en sus síntomas y en su fisiopatología, es decir, en el mecanismo por el que se produce esa “enfermedad”.

Hablamos de patología general porque nos interesa el estudio global que sienta las bases de cualquier posterior subdivisión analítica. Podríamos hablar también, *sensu strictu*, de patología social (término introducido por la escuela organicista) de la evaluación, ya que se trata de desórdenes de procesos que afectan a grupos y a relaciones entre individuos.

La evaluación (tanto la de carácter funcional como la investigadora) puede ser manejada para servir los intereses del evaluador ya que puede llamar "evaluación" a la operación que desee, puede evaluar aquello que le interese, en las formas y momentos que determine, con los instrumentos que considero oportunos y -desde luego- para utilizar en los fines que su particular interpretación aconseje.

El poder que dinámica del proceso evaluado es tan grande que permite atribuir causalidades arbitrariamente, catalogar éticamente, justificar decisiones y clarificar realidades...

La apariencia de rigor se convierte en una amenaza de perfiles contundentes. Los riberes científicos disminuyen la capacidad de crítica y prestan un peligroso aval. Si entendemos/practicamos la evaluación como un proceso difícilmente contestable en su concepción, en su desarrollo, en su utilización, difícilmente podrán ser discutidas las conclusiones: "Este alumno ha suspendido". "este centro escolar es malo", "esta experiencia educativa es excelente", "esta actividad tienen que desaparecer", "estas personas tienen que ser excluidas"...

Los criterios que se aplican para la evaluación no siempre se ajustan a patrones rigurosamente elaborados. Y así, una reforma puede considerarse "buena" porque aumenta el número de los conocimientos de los alumnos, un centro puede etiquetarse como "estupendo" porque aprueba la selectividad un porcentaje alto de alumnos presentados y un alumno es considerado "excelente" porque ha contestado correctamente una prueba objetiva.

Más aún, de una evaluación superficial pueden nacer explicaciones explícitas/implícitas que consagran una determinada realidad, estableciendo nexos causales gratuitos. "Este alumno suspenderá porque no estudia", "ese programa de renovación no funciona porque los profesores son incompetentes", "ese centro tienen eficacia educativa porque posee grandes instalaciones"...

La patología que afecta a la evaluación atañe a todas y cada una de sus vertientes. Por qué se evalúa (y para qué), quien evalúa, cómo se evalúa, para quién se evalúa, cuándo se evalúa, qué se evalúa, a quién se evalúa, con qué criterios se evalúa, cómo se evalúa la misma evaluación, etc.

Algunas patologías lo son en cuanto hipertrofian un aspecto o dimensión que, planteando en su justa medida, sería positivo. Valorar los conocimientos, por ejemplo, no es un error, pero sí lo puede ser el valorar exclusivamente, obsesivamente, los conocimientos, sin tener en cuenta su naturaleza, su importancia, su interés, su adecuación, su coordinación con otros conocimientos, etc.

Otras patologías se deben a la atrofia de funciones que son consubstanciales con un estado de salud educativa,. No se desarrolla, por ejemplo, un tipo de evaluación democrática en la que tanto el proceso de evaluación como el manejo de la información resultante es responsabilidad directa de los protagonistas de la experiencia educativa.

Hay también carencia, disfunciones...La gravedad de la anomalía dependerá, en parte, de su misma naturaleza y, desde luego, de su intensidad.

1. SOLO SE EVALÚA AL ALUMNO.

En este sentido sí es protagonista el alumno. Se le examina siguiendo una temporalización determinada, se le dan los resultados, prácticamente inapelables y, en general, se le considera el único responsable de los mismos.

No parece concebirse el curriculum sin la evaluación del alumno, pero sí sin la evaluación genérica del mismo. Los argumentos –cargados de lógica- que se utilizan para avalar la ineludible necesidad de la evaluación del alumno no se aplican a otros elementos del currículo.

A cada alumno se le asigna en el expediente un valor numérico (al menos, cuantificando) que parece ser de su exclusiva responsabilidad. La calificación del alumno – para muchos padres, profesores y para los mismos alumnos- es el resultado de su capacidad y su falta o derroche de esfuerzos. En el caso de fracasar será él quién deberá pagar las “consecuencia”. Sólo él deberá cambiar. Lo demás, podrá seguir como estaba, la evolución se convierte así en un proceso conservador.

La distinción socorrida de evaluación suficiente y evaluación satisfactoria vienen a poner de manifiesto otro flanco de la patología. Porque se utiliza como criterio referencial táctico la consecución de unos conocimientos mínimos (será el profesor quien lo compruebe según sus particularidades criterios), estableciendo en la gama de puntuaciones una comparación con el resto de los escolares a todas luces discutible.

¿Qué es lo que se ha v comparado para colocar a los individuos en la escalera de suspensos/aprobados/notables/sobresalientes/matrículas...? No se sabe sí se comparan las capacidades de los sujetos, los esfuerzos realizados, los conocimientos adquiridos..., o la suerte de que han gozado. Lo cierto es que la etiquetación que nace de la evaluación educativa figura estampada en impresos, debidamente rubricada por la autoridad académica y refrendada por los ellos oficiales.

Este ejercicio enmascara una injusticia grande, no sólo por la arbitrariedad de asignación, por el capricho atributivo (agravado por la apariencia de rigor

científico), sino por la desigualdad radical de condiciones naturales y contextuales. Partir de situaciones desiguales y pretender comparar los resultados utilizando los mismos raseros es una fórmula radicalmente injusta de ejercitar una aparente justicia.

Lo cierto es que, en este mecanismo, sólo se incluye al alumno, quiera o no, quedando muchos responsables del proceso educativo sin esa consideración evaluadora y sin las consecuencias que llevaría aparejadas.

2. SE EVALÚAN SOLAMENTE LOS RESULTADOS

Los resultados han de ser tenidos en cuenta dentro del proceso evaluador. Pero no solamente los resultados. Los presupuestos de los que se parte, las condiciones que se tienen, las estrategias que se ponen en marcha, los procesos que se desencadenan, los ritmos de consecución, la proporción rendimiento/ esfuerzo..., son también elementos que deben evaluarse. No sólo porque la consecución/ no consecución de unos resultados (y el grado de su logro) está supeditada a aquellos factores sino porque ellos mismos constituyen el objetivo de la mirada evaluadora. En definitiva, no sólo importa qué es lo que se ha conseguido, sino el cómo, a qué precio, con qué ritmo, con qué medios, con cuántos esfuerzos, a qué consta, para qué fines.

Analizar sólo los resultados obtenidos es, cuando menos, parcial. Y la parcialidad suele ir acompañada de imprecisión y de tergiversaciones.

3. SE EVALUARÁN SOLO LOS CONOCIMIENTOS

El proceso de enseñanza/aprendizaje se realiza sobre un cuerpo de conocimientos más o menos estructurados, más o menos interesantes, más o menos conexiones con la práctica, más o menos conexiones con la práctica, más o menos "autónomos" (los grados de libertad del currículum pueden ser viables). No se puede aprender en el vacío. Cuando hablamos de "aprender a aprender" – dejando al margen los conocimientos – estamos haciendo meras piruetas mentales.

Aprender a aprender es un slogan tan utilizado como desprovisto de sentido real. Porque sólo se aprende aprendiendo.

No se puede, rechazar el aprendizaje de contenidos, porque son necesarios para articular el pensamiento, para adaptarse a la realidad y para poder manejarla.

Otra cosa es la selección de los contenidos, su articulación, su significación de organizadores del pensamiento.

Ahora bien, limitarse a la evaluación de conocimientos supone un reduccionismo escandaloso. Existen otra serie de pretendidos logros (Véase cualquier formulación curricular, véase la legislación de cualquier rango que defina objetivos

educativos del sistema) que no se contemplan debidamente en el proceso evaluador: actitudes, destrezas, hábitos, valores...

Una persona que adquiere un abundante caudal de conocimientos para mejor destruir/oprimir a los otros, un alumno que llenase su cabeza de conocimientos pero que odiase la sabiduría, un individuo con gran almacén de datos en la cabeza pero incapaz de comunicarse..., no estaría auténticamente formado.

Luego veremos que la evaluación de estas facetas no es tarea fácil. Muchos profesores no tienen conciencia profesional de que se trata de aspectos educativos relevantes. Otros no conocen las formas de acercarse a una evaluación adecuada de los mismos. La administración, concedora de esas limitaciones nacida de una deficiente formación del profesorado, desconfiando de que puedan realizar espontáneamente estas tareas, pretende implantarlas a golpe total de "B.P.E." Es más barato que una larga y concienzuda formación. Pero es totalmente ineficaz. ¿De qué sirvió implantar la evaluación continua cuando los profesores seguían instalados en sus viejas prácticas evaluadas? Sencillamente, para repetir continuamente aquellas prácticas de evaluación memorística.

4. SOLO SE EVALÚAN LOS RESULTADOS DIRECTOS, PRETENDIDOS

La puesta en marcha de muchos proyectos, curriculares que siguen un criterio de encadenamiento lineal (objetivos propuestos. Contenidos-métodos – evaluación de objetivos propuestos) no tienen en cuenta la evaluación de aquellos efectos laterales, secundarios, imprevistos.

Estos efectos existen, casi inevitablemente. Y, en ocasiones son mucho más importantes que los resultados buscados directamente en el proyecto.

La pretensión de que el alumno adquiera un elevado nivel de conocimientos, ejercida de manera despótica, puede engendrar una aversión hacia el estudio muy perjudicial.

La pretensión, en algunos centros religiosos, de que los alumnos adquieran una práctica religiosa acendrada, llevó a muchos a una postura de rechazo y resentimiento hacia lo religioso.

Lo mismo habría de decirse de un centro o de una reforma. El resultado directo del programa es evaluado, pero sin tener en cuenta los efectos secundarios (llamémosles así, en ocasiones tengan mayor relevancia que los buscados) que se han generado.

Introducir una mejora en un centro puede suponer, quizás, el enfrentamiento entre dos sectores del claustro o en enrarecimiento de las relaciones entre la dirección y los profesores.

Bien es cierto que los efectos secundarios, en un momento incontrolados e impredecibles, no deben frenar la innovación o la implantación de un programa seriamente estudiado. Lo que decimos es que la evaluación del mismo debe tener en cuenta tanto los resultados que se buscan como los que se hayan provocado a lo largo del desarrollo curricular.

5. SOLO SE EVALÚAN LOS EFECTOS OBSERVABLES

Existen efectos que son directamente observables., estos efectos-buscados o no suelen pasar inadvertidos a los ojos del evaluador.

Los modelos de evaluación basados en la programación por objetivos operativos (Pager, 1972; Bloom, Popham, 1975), exigen un modelo correctivo de evolución que descansa sobre la aplicación de instrumentos de comprobación de carácter experimental.

Se podría objetar que no es posible evaluar los efectos no observables. No es así. Lo no observable no es equivalente a lo no existente. Ni a lo no relevante. Ni por supuesto que pretenda profundizar en la entraña educativa.

Eisner (1981) dice que la mayoría de los aprendizajes que realizan el alumno en la escuela no se hallan programados en el curriculum explícito.

Las limitaciones metodológicas que impone una concepción experimentalista de la evaluación, no debe hacernos olvidar todo el bloque de iceberg curricular que permanece por debajo de la superficie de las aguas.

Es precisamente un, modelo de evolución más exigente, más riguroso, más ambicioso el que, paradójicamente, nos lleva a explorar otras capas que no son reconocibles "a simple vista" o a golpe de mediaciones pretendidamente objetivas.

Estas dimensiones de la evolución exigen, eso sí, la utilización de técnicas de exploración adecuadas al intento de llegar a descubrir e interpretar lo culto del curriculum y de sus resultados.

6. Se evalúa principalmente la vertiente negativa.

En la práctica habitual del docente la evaluación está marcada por las correcciones. El mismo lenguaje descubre la actitud predominante: "corregir" significa "enmendar lo errado". El subrayado de las faltas de ortografía es mucho más frecuente que la explícita valoración de las palabras bien escritas.

Sirva esta anécdota de punto de referencia para referirnos a esta actitud no solamente mala por negativista sino por lo parcializante.

Una evaluación rigurosa requiere un tratamiento holístico de los fenómenos y de los productos. La comprensión de un proceso adquiere sentido en un análisis estructurado y estructurante en el que la interconexión de todos los elementos permite la explicación y el significado.

El desequilibrio de perspectiva hace que la escuela esté más atenta a los errores que a los aciertos de los alumnos y que los equipos de evaluación externa se apresten más a descubrir problemas y deficiencias que a resaltar valores y logros.

Además de factores actitudinales puede influir en esta dinámica selectiva la mayor facilidad que existen de describir la tensión que la calma, el error que el acierto, la guerra que la paz...

7. SOLO ASE EVALÚA A LAS PERSONAS

Es un error "someter" a los alumnos a los profesores de un Centro a los coordinadores de una reforma a una evaluación conclusiva (lo cual encierra juicios de valoración, no meras descripciones de actuación) sin tener en cuenta las condiciones, los medios, los tiempos, los contextos en que se mueven.

Cronbach (1963) distingue tres grandes áreas sobre las que la evaluación toma decisiones. El material de instrucción, los individuos y la regulación administrativa. No son sólo los individuos los responsables de un proceso o de un resultado. Hay que contemplar cuáles son los medios con los que cuentan, las condiciones en las que trabajan y los márgenes de autonomía real con que cuentan.

Los mecanismos de atribución pueden hacer descansar en el sistema, en la carencia de medios, en la presión de los programas oficiales, en la mala organización de los centros, en la irracionalidad de los calendarios..., toda la responsabilidad de un proceso educativo.

Esos condicionantes/estimulantes existen, hay que tenerlos en cuenta a la hora de la evaluación. No solamente como coadyuvantes del intento educativo, sino como elementos que puedan ser modificados y/o mejorados para un posterior desarrollo.

8. SE EVALÚA DESCONTEXTUALIZADAMENTE.

Benttanffy (1976, 1978) plantea la necesidad de tener en cuenta un contexto amplio cuando se pretende comprender la desliad de un sistema actuante.

Pretender dar significado a la actuación de un alumno desde la óptica y el código del evaluador, prescindiendo de las claves de interpretación del contexto, es vaciar el contenido de la realidad (Oates, 1975).

Encasillar un Centro de la plantilla elaborada por el evaluador con unos criterios genéricos de pretendida validez, es negarse a entender todo lo que sucede en el mismo.

Una calificación sobresaliente puede ser considerada "algo despreciable" en un contexto determinado. Una clase indisciplinaria puede estar en el eje de la administración de la mayoría de los alumnos de un Centro. Una experiencia pedagógica "modélica" puede ser valorada d en su contexto de forma negativa...

Si el evaluador, en aras de una pretendida objetividad, busca la creación y realiza la aplicación de instrumentos de métrica fiables y válidos técnicamente y no tienen en cuenta la realidad viva, compleja y dinámica de eso todo, de ese sistema organizativo que tienen en sí mismo los códigos sintácticos y semánticos, se verá enredado en un caudal de datos muertos y desprovistos de auténtico significado.

La actuación de un alumno podrá ser entendida e interpretada justamente (con justeza y con justicia) en el marco de innumerables redes codificadoras que se producen en el sistema de aula que, a su vez, está en conexión con el sistema del Centro que a su vez...

"La optimización de un sistema es posible, siempre y cuando nos acerquemos a las peculiaridades específicas de la estructura y comportamiento de cada sistema, y en función precisamente de ese conocimiento específico" (Pérez Gómez, 1985).

9. SE EVALÚA CUANTITATIVAMENTE.

La pretensión de atribuir números a realidades complejas es un fenómeno cargado de trampas en el área de la educación.

En las calificaciones escolares utilizamos escalas nominales, ordinales y de razón. Un aprobado es distinto de un suspenso, un 5 es una nota inferior a un 8, un 6 es el doble de un 3. Parece que todo está claro, que todo es muy preciso.

El peligro de la evaluación cuantitativa no es solamente la imprecisión sino –y sobre todo- la apariencia de rigor. "la asignación de números de una manera mecánica, como es común en los procedimientos cuantitativos, no garantiza la objetividad" (Cook, 1986).

Pero como aparentemente tienen objetividad, genera en los usuarios y destinatarios una tranquilidad mayor que mata las preguntas más hondas.

Algunos padres de alumnos que reciben informes sobre la marcha de sus hijos con la expresión "Progresan adecuadamente", preguntan a sus profesores: "Progresan adecuadamente, ¿equivale a un 8 de antes?". La puntuación parece un

lenguaje más claro, engañosamente más claro-. El padre tienen, además, el punto de referencia de otros alumnos que tienen puntuaciones de 4, 6, ó 9.

También el profesor puede sentirse más "seguro después de estampar sobre la prueba objetiva, corregida con una fórmula matemática ($P = A - E / N - 1$), el número exacto, con decimales, correspondientes a la calificación. Además puede establecer claramente, matemáticamente, la línea del apto/no apto.

El alumno, con este procedimiento calificador, sabe lo que tiene que estudiar, cómo ha estudiarlo y –después de la calificación- saber cuánto ha aprendido.

El problema, además del que entraña la pretendida objetividad de esa puntuación, estriba en que la luz cegadora de esa "claridad" no deja ver cuestiones más importantes:

- ¿Cómo aprende el alumno?
- ¿Cómo relaciona lo aprendido?
- ¿Cómo inserta los nuevos conocimientos en los ya asimilados?
- ¿Ha disfrutado aprendiendo?
- ¿Estudiaría "esas cosas" por su cuenta?
- ¿Tienen ganas de aprender cuando terminan las pruebas?
- Etcétera.

10. SE UTILIZAN INSTRUMENTOS INADECUADOS.

En una recopilación de instrumentos en curso p para la evolución de centros escolares (Plaza, 1986), se ha podido comprobar que la casi totalidad tienen una configuración estática, anecdótica, aséptica, cuantificadora, descontextualizada, etc.

Los cuestionarios, escalas de estimación, listas de control, plantean el interrogante de la triple adecuación realidad-percepción, percepción-interpretación, interpretación-expresión. Se añade a esta problemática la complejidad de funcionamiento de la biblioteca del Centro nos deja en la duda de si la baja puntuación se debe a algunas causas (o a otras posibles):

- No hay libros.
- No está actualizada la documentación.
- No existen ficheros que permitan la localización rápida y fácil.
- No existen locales.
- No hay tiempo para su apertura.
- No hay afición por la lectura.
- Los libros no tienen interés para los alumnos.
- No están formulados los alumnos de sus normas de funcionamiento.
- Existe mucho ruido y es difícil la concentración.
- El bibliotecario es una persona indeseable.

- Está mal visto entre los alumnos acudir a la biblioteca.
- Etc.

Decir que el 48% de los profesores valoran con una, medida de 3 puntos el funcionamiento de la biblioteca, es un modo de esconderse, bajo la cifra porcentual, las innumerables causas que han llevado a los opinantes a conceder una puntuación.

Ese motivo de evaluar hace difícil la comprensión de la realidad, su interpretación y, obviamente, la posibilidad de efectuar que mejoras que solucionan los problemas o potencien los aciertos.

Lo mismo sucede en el aula. Un instrumento de valoración "objetivo" no puede estar más cargado de subjetividad/arbitrariedad. Ha aquí una escala de ambigüedad que no siempre es tenida en cuenta:

- ¿Lo que aparece en el curriculum como contenido mínimo es lo realmente importante, valiosos, interesante...?
- ¿Lo que ha seleccionado el profesor para la prueba es significativo de lo que tenía que aprender?
- ¿Lo que pregunta el profesor es exactamente lo que quiere saber sobre lo que ha aprendido el alumno?
- ¿Lo que el alumno lee es lo que el profesor ha querido preguntar?
- ¿Lo que responde es exactamente aquello que sabe la cuestión?
- ¿Lo que interprete el profesor es lo que el alumno realmente ha expresado?
- ¿La valoración corresponde a lo que el profesor entiende que ha expresado el alumno...?

Nada digamos de la dificultad de mejorar el aprendizaje a través de ese tipo de pruebas. A lo sumo, el alumno prestará atención al mejor modo de contestarlas en próximas ocasiones. Lo cual es distinto al propio proceso de enseñanza/aprendizaje.

Existe otra complicada cuestión en la aplicación habitual de este tipo de pruebas "objetivas" bajo el pretexto de que son más justas, ya que "miden" a todos por igual. Lo cierto es que no existe arbitrariedad que la de querer "medir" de la misma forma a personas que son radicalmente diferentes. Así, un alumno tímido no deseará realizar una entrevista o examen oral. Quién no se expresa bien por escrito, preferirá una prueba...

La naturaleza de la prueba lleva dentro de si un componente sesgado de valoración, independientemente del tipo de contenidos que plantee y de su forma de presentarlos.

11. SE EVALÚA DE FORMA INCOHERENTE CON EL PROCESO DE ENSEÑANZA/APRENDIZAJE.

El camino es circular, no rectilíneo y unidireccional. Se valúa como se enseña y se enseña como se evalúa. O más bien, se estudia para la evaluación. De tal forma que es este proceso el que dirige el aprendizaje. Los alumnos estudian para el examen. No sólo en función de ese momento sino "de forma" que les permita hacer frente al mismo con suficientes garantías de éxito.

Si el examen consiste en una prueba objetiva de Verdadero-Falso estudiará u de forma distinta que si se le anuncia una prueba de ensayo o si se le exige un diseño creativo. Es más, los alumnos tratarán de acomodarse a las expectativas del profesor, a sus códigos de valor.

La incoherencia se establece cuando se quiere realizar un aprendizaje por comprensión y se realiza luego una prueba de carácter memorístico, rígido y repetitivo.

Un proceso de enseñanza basado en la explicación oral se cierra con un modelo de examen escrito.

Un modo de trabajo asentado sobre el grupo concluye en una evaluación individual.

Un proceso de enseñanza/aprendizaje teóricamente asentado sobre el desarrollo integral del individuo, acaba con una evaluación exclusivamente preocupada por los conocimientos adquiridos.

Otras veces, el enfoque exclusivo en el aprendizaje de contenidos intelectuales pretende concluirse con una evaluación que va mucho más allá de lo que se ha trabajado.

En el área de la didáctica se puede rizar el rizo al exigir el profesor la realización de una prueba objetiva por él diseñada sobre el tema: "Importancia, necesidad y valor del proceso de evaluación en la enseñanza".

12. Se evalúa competitivamente

Creo que una de las ventajas que encierra el enfoque cualitativo es el afinar la sensibilidad del evaluador ante los procesos.

El auténtico significado de un proceso educativo, se encuentra en el análisis de todos los elementos que lo integran y que lo circunscriben. Pero no en e con—traste, comparación y competencia con otros procesos similares, diferentes o antagónicos.

La evaluación habitual del alumno parece cobrar sentido en la comparación y la competencia. Para algunos padres, profesores y alumno importa el cuándo del aprendizaje. Y que ese cuándo esté por encima de...

El carácter positivo que se pretende dar a la evaluación en algunos ámbitos educativos, se trastoca frecuentemente en una burda o camuflada contienda. "Saber más que", "saber menos que", son puntos de referencia cruciales.

Lo mismo sucede con la evaluación de Centros. La valoración cuantitativa es propensa a la comparación. Las escalas tienen en ella su justificación más evidente. Este es un Centro "mejor" porque ha obtenido en la evolución más puntos que otro.

La localización de la mirada se sitúa en el exterior del proceso mismo, con el agravante de que se establecen comparaciones arbitrarias ya que pocas veces se para de condiciones, medios, historia, personas y fines similares.

Por otra parte, es frecuente justificar la comparación con el pretexto de que sirva de estímulo y emulación. Durante toda la enseñanza primaria y secundaria tuve que trabajar con émulos. Cada alumno tenía un compañero que era un competido. Con él había que medirse. El éxito consistía en aventajarle.

Se desvía la atención del auténtico proceso de aprendizaje de cada individuo, de cada aula, de cada Centro. Y si es posible pensar que el "ganador" se sentirá satisfecho y el "perdedor" estimulado, no es menos imaginable que el primero se sienta ridícula y estúpidamente orgulloso, y el segundo humillado.

En las evaluaciones que he visto realizar sobre los profesores (por sus alumnos), siempre ha visto solicitar y exigir cautelas para que los resultados no se hiciesen públicos a través de escalas valorativas. "No se trata de establecer comparaciones", se dice.. los profesores no consideran positivo que se publiquen unas listas donde unos aparecen como los mejores y otros los peores. Pero colocan, sin ambages, las calificaciones escalonadas de sus alumnos, muy bien matizadas con decimales...

13. SE EVALÚA ESTEREOTIPADAMENTE

Los profesores repiten una y otra vez sus esquemas de evolución. Cada año los alumnos se preocupan de saber cuál es la costumbre evalaudora del profesor.

De forma casi automática., el profesor repite sus fórmulas. Ni siquiera negocia con los alumnos el planteamiento habitualmente practicado. Al comienzo de curso fija el número, la forma y los matices.

¿Ha sometido a evaluación sus propios mecanismos de evolución? No de una forma rigurosa. Cuando ha pensado en ello ha sido para confirmar los estereotipos.

En un curso con cinco asignaturas un alumno deberá "someterse" a cinco proyectos diferentes de evaluación. Los profesores evalúan de formas muy diferentes. Pero cada profesor, eran los diferentes cursos en que impartía docencia, evalúa de una forma idéntica.

Se multiplican los trabajos sobre evaluación pero no son los protagonistas -los propios profesores- los que emprenden esta labor. Stenhopuse (1984) dice que lo que mejora la acción educativa no es tanto la investigación sobre los profesores sino la de los profesores que analizan su propia actividad.

En los Centros, la mecánica de la confección de las Memorias tiende a convertirlas en pura rutina. Y, en la medida que se establezcan pautas "generalizables", estas rutinas serán similares en casi todos los Centros. Tan similares como inútiles.

14. NO SE EVALÚA ÉTICAMENTE

Además de los problemas técnicos acechan al proceso evaluador numerosos conflictos de carácter ético.

La evaluación puede convertirse en un instrumento de opresión. ¿Qué sucedería en las aulas si el profesor estuviese desprovisto del "arma" de la evaluación?

Cundo se articula el proceso de enseñanza/aprendizaje sobre el resultado de la evaluación -más que sobre la riqueza y la profundidad del saber- se corre el riesgo de la manipulación y el sometimiento del alumno,. La "hora de la verdad" no es la del aprendizaje sino la de la evolución. Cuando es el profesor quien lo decide todo respecto a este momento decisivo, todo el poder descansa en sus manos.

La evaluación ha sido un instrumento de control, de amenaza e, incluso, de venganza, respecto a algunos alumnos que se han permitido ejercitar el derecho a la crítica, a la discrepancia o la indisciplina.

Cuando apoyándose en la evaluación, se establecen nexos causales poco rigurosos, habrá que pensar cuáles son las intenciones y los intereses que motivan la manipulación.

En la evaluación institucional se plantean numerosos problemas de carácter ético,. Uno de los es la confidencialidad de los informes. ¿De quién son los datos? ¿Para quién son los daros? No puede revelarse el contenido de los informes identificables, tanto por lo que respecta al emisor como al destinatario de los juicios.

La independencia de los equipos que realizan evaluación externa al servicio de la Administración o de entidades que financian los proyectos, no siempre está garantizada. John Elliott, Helen Simona y Savel Kousner insisten (Elliott et al., 1986) en la necesidad de conseguir la independencia de los técnicos frente a las posibles exigencias del poder (académico, político, financiero, etc.).

A nadie se le oculta que una evaluación puede ser rígida, condicionada o manipulada al servicio de unos intereses determinados, o de unas decisiones pretendidas: retirar subvenciones, modificar proyectos, sustituir personal, clausurar centros, cortar experiencias, etc.

15. SE EVALÚA PARA CONTROLAR

La evaluación en educación, paradójicamente, no suele ser educativa. O repercute en la mejora del proceso. La evaluación se cierra sobre sí misma, constituye un punto final.

Se habla de "calificación final". Cuando es justamente un momento más en el momento de la mejora. No se debe confundir control con evaluación, aunque las constituye su esencia más rica, más dinámica.

Cuando los profesores se niegan a explicar a sus alumnos (o se muestran reticentes a hacerlo, por pensar que se trata una pérdida de tiempo) de dónde proceden las calificaciones que les ha atribuido, están desaprovechando un buen elemento de aprendizaje.

Disparar con los ojos vendados sobre una diana sin saber dónde se ha producido el impacto,, hace imposible la mejora de la puntería en ensayos posteriores. Solamente conociendo el resultado se puede mejorar en próximos disparos. Pero, no solamente será preciso conocer y analizar lo que ha sucedido, sino planificar los nuevos procesos en función de aquello que se ha descubierto como fracaso o acierto.

Ese efecto retroalimentador se suele perder en las evaluaciones educativas. Y así, la evaluación que suponen las memorias finales de los Centros, en nada iluminan el nuevo proyecto en el curso siguiente, los puentes de la eficacia están valorados por la falta de análisis, la inercia inveterada y la despreocupación de muchos profesionistas.

El mundo educativo esta lleno de situaciones en las que no se aprovecha esta riqueza potencial que lleva en su interior el proceso evaluador. Los profesores/alumnos repiten los mismos errores casi no obstinados, los cursos que finalizan una especialidad en nada benefician con el análisis de su historia a los

cursos siguientes, los Centros no aprenden de la reflexión rigurosa sobre su experiencia...

"La evaluación es el proceso de diseñar, obtener y proporcionar información útil para juzgar alternativas de decisión", dice Stufflebeam (1971). Según este planteamiento, cuando la información no repercute en la toma de decisiones perdería su misma esencia.

La nueva decisión, al ser puesta en práctica, deberá ser también evaluada. Este proceso dinámico, abierto, facilita el cambio y posibilita la mejora.

16. SE EVALÚA PARA CONSERVAR.

Hemos apuntado ya que la evaluación suele cerrarse sobre sí misma, limitándose a desarrollar una función sancionadora. En este sentido, no impulsa el cambio.

Pero, hay algo más grave, en escondido en la patología de su funcionalidad. La utilización de la misma para justificar el status quo del profesor, de la dirección del Centro, de la Administración Escolar.

Si los malos resultados o el mal funcionamiento del proceso se pueden atribuir libremente (arbitrariamente) a quien se quiera con la apariencia de la objetividad, del rigor científico y del argumento de autoridad del que maneja la evaluación, el "interesado en no cambiar" que, a lo postre, tienen en su poder la interpretación de la evaluación, podrá mantener tranquilamente los mismos planteamientos.

Y los cambios introducidos afectarán a las capas y esferas menos comprometedoras, más superficiales o más económicas a la hora de realizar el diagnóstico de la innovación y la estrategia y la temporalización de la misma.

Lo cierto es que se evalúa constantemente en la escuela, pero se cambia muy poco.

17. SE EVALÚA UNIDIRECCIONALMENTE

La evaluación tienen sentido "descendente". El Ministerio "evalúa" a inspectores, los inspectores "evalúan" a los directoras escolares, los directores escolares "evalúan" a los profesores, los profesores "evalúan" a los alumnos. Si bien es cierto, que en la parte más baja es donde se nota el peso de esa función.

No se produce una evaluación en sentido "ascendente". No, desde luego, en sentido horizontal.

Un modelo de evaluación democrática supone que los interesados manejan la evaluación. Deciden sobre ella. Y ellos son los que dicen lo que piensan, los que analizan lo que hacen.

“Para conocer la realidad de sus significados relevantes hay que sumergirse en el curso real y vivo de los acontecimientos y conocer las interpretaciones diversas que hacen de las mimas aquellos que las viven” (Pérez Gómez, 1983^a).

Sólo la evaluación democrática puede propiciar un cambio en profundidad . Esta evaluación, solicitada o emprendida por la comunidad en cuestión, revierte sobre ella en informes rigurosos, claros y significativos.

No es que la evaluación democrática pueda prescindir de los expertos. Pero no son éstos los que puedan atribuir significado de espaldas e al mundo de los protagonistas. Y no son los expertos los que tienen el poder de la evaluación. Los expertos tienen la técnica, pero no las claves de la interpretación ni los resortes del poder pedagógico.

En la evaluación externa sólo la participación directa de los protagonistas podrá dar fiabilidad y validez a la lectura de los signos, a su metalectura posterior y a su utilización para el cambio.

Los que desean conocer el resultado de la evaluación y los que la llevan a cabo son integrantes de un proceso. También por eso la evaluación debe ser democrática (McDonald, 1976).

18. NO SE EVALÚA DESDE FUERA

Una experiencia educativa necesita la evaluación externa para poder realizar una mejora sustantiva. No hacerlo así significa cerrar el horizonte valorativo y arriesgarse a la miopía y a la deformación óptica del que mira algo que está sobre sus mismos ojos.

Cerrar el círculo de la planificación de objetivos, métodos, evaluación, desde la exclusiva mirada del protagonista del proceso entraña el riesgo de quedarse atrapado en la propia limitación.

El evaluador externo goza de un punto de vista privilegiado, en cuanto tiene:

- Una distancia afectiva de la dinámica y del resultado.
- Unos criterios de independencia respecto al resultado.
- Unos puntos de referencia más amplios y complejos.
- Una mayor disponibilidad en el tiempo y la dedicación.

Esto no quiere decir que se pueda prescindir de los protagonistas de la experiencia, como apuntábamos más arriba. Tanto es así, que dentro de la metodología etnográfica de la evaluación es aconsejable la aportación de observador *participante*.

No decimos que el evaluador externo tenga que ser aséptico o ejercer una neutralidad que no es posible ni siquiera deseable. El evaluador no solamente mira, sino que busca. Y para ello ha de interpretar. Y la interpretación necesita unos códigos estructurados en los ejes de una sólida teoría.

19. NO SE HACE AUTOEVALUACIÓN.

La autoevaluación es un proceso de autocritica que genera unos hábitos Enriquecedores de reflexión sobre la propia realidad.

Dice Popper que realizamos más progresos al reflexionar sobre nuestros errores que al descansar en nuestras virtudes.

John Elliot plantea la autoevaluación de los Centros como una tarea de investigación en la acción (1986).

Las técnicas de diseño, de explotación y de interpretación y utilización de esas autoevaluaciones han de ser conocidas, asumidas y llevadas a término por los miembros de la comunidad escolar. ¿Tendría sentido un proyecto de evaluación que no fuese deseado, aceptado y conocido por posprotagonistas de la experiencia?

Muchas de las reticencias, de los miedos y de los conflictos que provoca una evaluación externa/impuestas desaparecen, ipso facto en el proceso de evaluación.

También los alumnos pueden/deben practicar estos procesos autoevaluadores. El profesor debe. El profesor debe poner en sus manos los instrumentos precisos para ello.

Y han de negociar con ellos el reparto de las cotas de decisión que lleve consigo la evaluación.

No es aceptable la práctica del juego autoevaluador, en la cual el alumno hace reflexiones y análisis sobre su aprendizaje, pero no puede materializar el resultado en una parcela de las calificaciones.

El profesor suele actuar con muchas reticencias en este sentido, suponiendo que el alumno son se calificará con criterios justos, por falta de objetividad o por carencia de referencias exteriores que le sirvan de contraste.

Contaré una experiencia reveladora al respecto, en un claustro de profesores de un Centro Escolar oí cómo los profesores se resistían a practicar la autoevaluación con sus alumnos por considerar que eran "demasiado pequeños, que no tenían capacidad parao ello". Ese mismo día en un Departamento universitario, otros profesores argumentaban que "la autoevaluación de los alumnos no era posible en el ámbito universitario porque no tenían experiencias de haberla practicado en Centros de niveles inferiores".

Cuando los profesores utilizan estas excusas no están en condiciones de realizar esa experiencia. Del mismo modo que cuando dicen que el número de alumnos o las condiciones de trabajo y de tiempo les obligan a realizar una evaluación superficial y masiva, no podemos afirmar con seguridad que con menor número y mejores condiciones harían las cosas de otro modo.

20. SE EVALÚA "DESTEMPORADAMENTE"

Entiéndase aquí el prefijo en su acepción griega (dis=mal), a sabiendas de que la palabra compuesta no debería utilizarse.

Cuando se habla de "evaluación continua" se quiere decir algo que casi nunca se hace. Más bien ha tenido el efecto contrario como apuntábamos anteriormente. El deseo de eliminar una mala forma de evaluar (anecdótica, memorística, superficial...) llevó a multiplicar las veces en que esto se hacía.

Porque el núcleo del concepto está más en el cómo que en el cuándo, más en el fondo que en la forma. El concepto exigía una actitud distinta, u otros métodos diferentes, un enfoque nuevo. Pero no hacer muchos más controles del mismo tipo que los que se hacían.

No se hace, pues, una adecuada evaluación sincrónica respecto al proceso de aprendizaje o, más ampliamente, de educación.

Pero tampoco se hace una adecuada evaluación diacrónica,. Entendemos por "evaluación diacrónica" aquella que se realiza a través del tiempo, con una perspectiva temporal que ofrezca nuevos elementos de referencia.

Veamos un ejemplo relativo a la evaluación de Centros Escolares. ¿Cómo evalúan uno, dos, tres, cinco años? El "antiguo alumno" tendrá unos elementos de comprobación/constatación que no tienen el individuo que está inmerso en la experiencia. Sabe si "aquello" sirvió para algo, si le permitió enfrentarse a los problemas...

Además de estas ventajas se encierra la de una independencia más clara, más firme, más abierta. Ahora puedo opinar sin ningún tipo de reticencias.

En este grupo que llamamos ex -alumnos de un centro habrá que integrar a los que lo abandonaron y a los que -de una u otra forma- se les instó a cambiar de centro o a quedarse en la calle.

21. NO SE HACE PARAEVALUACIÓN

La evaluación a la que aquí nos referimos se ciñe a cuestiones que van más allá de la simple comprobación.

Veámoslo con algunos ejemplos. Un evaluador externo analiza la eficacia de una institución militar en la forma de sus alumnos. Y puede llegar a la comprobación de que realmente existe un buen planteamiento en los métodos que intentan lograr

esos objetivos. Realmente se ya conseguido. Supongamos que esos objetivos" no son deseables", que esos métodos utilizados "no son éticos" y que esos logros "son contraproducentes para la formación de la persona".

¿Termina la función evaluadora con el informe de que el funcionamiento es bueno porque se está desarrollando el programa de forma eficaz?

Un sistema educativo de un país totalitario incluye entre sus objetivos la formación de los profesores y de los alumnos en la filosofía y en el proselitismo sectario. Y el análisis del funcionamiento del sistema, permite al evaluar comprobar que se están consiguiendo esos objetivos de forma rápida, eficaz y profunda. ¿Termina su tarea al realizar un informe en el que se describa y analice la eficacia del sistema?

Entendemos que no. De alguna manera Scriven (1973) se refiere a esta cuestión cuando diferencia la evaluación de la estimación del logro de los objetivos.

Nosotros pensamos que la "paraevaluación" (para=junto a, al lado) supone un análisis de contenidos y un juicio de valor que va más allá de la simple descripción y análisis de la coherencia del programa y de la eficacia del mismo.

22. NO SE HACE META-EVALUACIÓN.

El proceso de evaluación es tan complejo que ha de ser necesariamente evaluador para poder atribuirle un valor.

En cualquiera de sus vertientes puede encerrar numerosas trampas, numerosos riesgos, numerosas deficiencias. Por eso se hace imprescindible establecer criterios que permitan evaluar los mecanismos de evaluación.

No solamente para analizar el rigor del proceso, sino para ver cómo esa función condiciona lo que se estaba haciendo. Ahora, la evaluación de un programa se convierte poderosa que puede distorsionarlo p modificarlo sustancialmente.

En e3sta vuelta de tornillo, vuelven a aparecer todos los componentes de la evaluación con un nuevo nivel de complejidad: quién metaevalúa, para quién, por qué, cuándo, con qué instrumento, qué, de qué modo.

Un proceso riguroso de metaevaluación no sólo permitirá valorar de manera rigurosa los resultados, sino que p´ermitirá decisiones eficaces p´ara mejorar el planteamiento, la dinámica y los modelos de evaluación.

ROSALES, Carlos; “Evaluación de los aprendizajes” En: Criterios para una evaluación formativa. España. Ed. Narcea pp 107-136

5. EVALUACIÓN DE LOS APRENDIZAJES

Al centrarnos en la evaluación de los aprendizajes, nos situamos en el eje del procesos didáctico, en el principal foco de referencia, de acuerdo con el cual van a resultar indirectamente evaluados todos los demás componentes y fases de la instrucción. Es necesario aquí volver a recordar los conceptos de evaluación diagnóstica, formativa y sumamente que habíamos mencionado al principio.

La simple evaluación sumativa, de carácter final, ha de ser progresivamente acompañada de los otros tipos de evaluación diagnóstica y formativa. La primera como punto de referencia para la comparación de los resultados del proceso didáctico con el estado inicial anterior al aprendizaje de una determinada unidad, o también, como indicador de las causas de un determinado fallo o retraso cuando éste se produzca. La segunda, evaluación formativa, como medio de perfeccionamiento «sobre la marcha» de la actividad didáctica en el momento en que se está produciendo y no ya cuando se realizó totalmente. Nos vamos a centrar ahora en distintos tipos de pruebas para la evaluación de los aprendizajes. Lógicamente, las características de los mismos van a variar según el tipo de aprendizaje a que se refieran. No serán igualmente aplicables las mismas pruebas para la evaluación del aprendizaje motor, intelectual o afectivo y aun dentro de cada uno de estos grandes ámbitos se pueden establecer diferencias. Una de las primeras características a tener en cuenta en las pruebas de evaluación es, en efecto, que cuenten con preguntas, cuestiones o ítems concernientes a los objetivos que previamente se hayan programado en la unidad objeto de evaluación. Otra importante cuestión a considerar es la adaptación de las mismas a la metodología utilizada. Si en un proceso instructivo se han previsto predominantemente objetivos intelectuales, las pruebas deberán constar de cuestiones sobre todo de este tipo. Si en la misma se ha desarrollado una metodología fundamentalmente teórica, no tendría sentido tratar de evaluar los aprendizajes con pruebas predominantemente prácticas. Otro tanto podría decirse de los recursos utilizados y del resto de factores que concurren en la instrucción. Todo ello podría encuadrarse genéricamente en lo que se suele denominar validez de las pruebas de evaluación.

Otra serie de aspectos a considerar son los que, también de forma global, denominaríamos condiciones de fiabilidad, que se refieren a la consistencia de los datos evaluados. Dicha consistencia se determina mediante la construcción de varias pruebas para un mismo aprendizaje o bien la utilización de la misma prueba en varias ocasiones diferentes con un intervalo de tiempo entre las mismas.

Finalmente, decir, en estas consideraciones preliminares que, ante la posibilidad de elegir una orientación evaluadora ante la referencia a la norma o la utilización de criterios, nos vamos a centrar en esta última, en congruencia con lo expuesto al principio sobre la evaluación formativa como instrumento encaminado a estimular el aprendizaje total de los

contenidos y el cumplimiento de todos los objetivos de cada unidad en línea con la teoría del «aprendizaje para el dominio».

Evaluación de aprendizajes del ámbito intelectual

A efectos puramente formales, de estudio, vamos a partir del supuesto de la existencia de tres grandes ámbitos de conducta y por lo tanto de aprender: intelectual, afectivo y psicomotor. En realidad resulta difícil, por no decir imposible, separar tajantemente aprendizajes pertenecientes a cada uno de ellos. No obstante, desde la perspectiva de predominio de uno de dichos componentes, nos aventuramos a separarlos sólo teóricamente. Según Gagné, en el ámbito intelectual se incluyen los aprendizajes de: información, habilidades intelectuales (conceptos, discriminaciones y reglas) y las estrategias cognoscitivas.

El aprendizaje de información abarca el conocimiento de hechos específicos como nombres de personas, lugares, elementos, etc.; el conocimiento de fenómenos o sucesos en los que se da ya una relación coherente entre hechos específicos y finalmente la organización de tales sucesos y hechos en relación con el contenido mental anterior ya poseído por el sujeto. La información puede ser obtenida por los más variados procedimientos: a través de la palabra oral o escrita, del lenguaje icónico o activo; a través de la palabra del profesor o la consulta de textos, etc. la actividad que nos muestra cuándo el alumno ha logrado el aprendizaje de información, es la de enunciar. Cuando el alumno enuncia por escrito u oralmente un determinado contenido de información, de manera coherente y significativa, podemos evaluar como positiva la adquisición por el mismo de la capacidad de información. Por ello, la actividad de recogida de datos para la evaluación, ha de centrarse en la conducta de enunciar. Si se utiliza la observación, ésta se ejercerá sobre tal conducta. Si se utilizan pruebas formales, éstas han de provocar en el alumno la actividad enunciativa. Por ejemplo: «Relata con tus propias palabras los sucesos ocurridos en tal fecha».

Entre las habilidades intelectuales incluye Gagné las de discriminar, conceptuar y aplicar reglas. Los conceptos pueden ser concretos y abstractos y las reglas simples y superiores. La discriminación faculta para distinguir entre las características de los objetos y de los símbolos que utilizamos. Ya antes de iniciar su período de escolaridad, el niño discrimina entre formas, tamaños y colores, etc. de los objetos, y luego, lo hará entre las características de los signos gráficos y orales del idioma materno. Cuando aprenda un nuevo idioma, lo hará con los signos de éste. La conducta indicativa de este tipo de actividad, de aprendizajes, se expresa precisamente mediante verbos como «discriminar» o «distinguir». Luego, la evaluación de esta habilidad se basa en la información que se pueda obtener (mediante observación o pruebas formales) sobre la actividad discriminatoria del niño. Como ejemplos de pruebas formales para la detección de aprendizajes discriminatorios podrían ponerse:

- En el periodo preescolar, que separe los bloques lógicos u otro tipo de objetos, distribuyéndolos en grupos según forma, color tamaño, etc.
- En el aprendizaje de un idioma que señale claramente la diferencia en la pronunciación de varios fonemas.
- En la enseñanza de las ciencias que sepa diferenciar diversos tipos de vegetación, etc.

La habilidad para conceptuar se adquiere sobre la base de la anterior discriminación y consiste, fundamentalmente, en la capacidad del alumno para establecer clases de objetos, fenómenos, seres... tras la previa discriminación de las características de los mismos. Puede tratarse de conceptos concretos o abstractos. En el primer caso asociamos un concepto directamente con la realidad y en el segundo somos capaces de asociarlo con una hipótesis verbal sin necesidad de referencia tan directa a la realidad. Las conductas que nos pueden mostrar la posesión por el alumno de tal habilidad son muy variadas. Para conceptos concretos se pueden señalar actividades como elegir, señalar, marca, pintar, etc. Así por ejemplo, podrían servir como cuestiones de una prueba evaluadora de conceptos concretos:

- Marcar con una cruz los objetos de forma rectangular
- Pintar de color rojo todos los dibujos que representan animales.
- Señalar varios objetos de forma alargada.
- Dibujar un círculo alrededor de cada muñeco.

En la evaluación de conceptos definidos o abstractos, debemos determinar si el alumno es capaz de hacer referencia no ya a objetos concretos sino a proposiciones o hipótesis verbales, etc. Así, «será capaz de asociar la palabra perímetro», con la proposición «conjunto de los bordes externos de una figura», en lugar de tener que establecer la relación directa con objetos de la realidad. Una clase de cuestiones para la evaluación de conceptos definidos podría ser precisamente esta actividad de asociación entre conceptos expresados mediante una palabra y las frases o posiciones explicativas de cada uno de ellos.

La habilidad de aplicación de reglas supone la capacidad en el alumno para la realización de una serie ordenada de pasos conducentes a la consecución de un objetivo determinado. Puede tratarse de reglas simples o de carácter superior, en las que se da una combinación de las primeras. En todo caso, es necesario no confundir el aprendizaje auténtico de reglas, basado en la comprensión causal de las operaciones, frente a una asimilación puramente verbal del enunciado de las mismas. Conductas significativas de estos tipos de aprendizaje serían las manifestadas mediante verbos como «demuestra» y «resuelve», que serán objeto de observación o provocación mediante cuestiones formales como:

- «Resolver una ecuación de segundo grado mediante la realización ordenada de los pasos incluidos en la misma».
- «Demuestra que conoce y sabe realizar las operaciones necesarias para cambiar una rueda a su automóvil».
- «Demuestra cómo se realizan diversas combinaciones de elementos químicos».

Las estrategias cognoscitivas constituyen un tipo de aprendizaje en virtud del cual el sujeto-alumno controla sus propios procesos mentales y cada una de las fases de los mismos en la adquisición del saber. La característica determinante de este aprendizaje por lo tanto, su proyección interna, sobre la propia actividad mental y no sobre la realidad exterior como ocurría con las habilidades. Existen estrategias correspondientes a cada una de las fases del aprendizaje: comprensión, adquisición, retención, recuperación, etc. Las estrategias se forman a lo largo de la actividad de aprendizaje y por esto son menos numerosas en los primeros años de escolaridad, en los que predomina el aprendizaje guiado y más abundantemente a medida que transcurren los años. Una conducta representativa de

este tipo de aprendizaje podría expresarse mediante el verbo «elabora». Así, se pueden considerar como cuestiones para la evaluación:

- «E alumno elabora un proyecto de viaje»
- «El alumno elabora un estudio climatológico»

Como ya se ha expresado al hablar de objetivos, Bloom ha clasificado los aprendizajes o conductas susceptibles de aprender en este ámbito, en las siguientes categorías: conocimiento, comprensión, aplicación, análisis, síntesis y evaluación. Las pruebas relativas a la evaluación en cada una de estas categorías presentan unas características propias, diferentes respecto a las demás, que según el auto deben ser:

Conocimiento

Las cuestiones destinadas a la evaluación de conocimientos, generalmente las más frecuentes en pruebas de evaluación, han de atenerse a dos *características* principales:

1. Presentar un *nivel de dificultad semejante* al que se ha desarrollado en la enseñanza correspondiente. No se debe exigir mayor exactitud o discriminación en las cuestiones de evaluación, pues en tal caso se estaría examinando al alumno de algo que no realizó, que no aprendió previamente.
2. Las cuestiones o pruebas han de presentarse en un *contexto similar* a aquel en que se realizó la enseñanza. Se refiere esta condición a la utilización de términos, materiales, situaciones análogas.

Dichas condiciones podrían hacernos creer que a través de esta forma de evaluación fomentásemos sólo la repetición, el memorismo. Esto sería realmente así en el caso de que evaluásemos sólo conocimientos. Por el contrario, éstos no son más que una categoría del conjunto de aprendizajes intelectuales a evaluar. Por ello, al lado de cuestiones para la evaluación del conocimiento, debemos incluir otras para la evaluación de la comprensión, la aplicación, etc., que poseerán ya características distintas como veremos

Tipos de cuestiones para la evaluación de conocimientos que se pueden utilizar:

—Pruebas de complementación:

«La guerra de la independencia duró... años».

—Pruebas de respuesta abierta:

«¿Qué es un círculo polar?»

—Pruebas de elección:

«¿Cuáles de las siguientes figuras son triángulos?»

Las pruebas de complementación están específicamente recomendadas en la evaluación del conocimiento de términos específicos, como nombres, fechas, lugares, fenómenos, etc. Se puede procurar evitar ambigüedades en su redacción. En ellas se valora fundamentalmente, pues, la capacidad de recuperación de información previamente asimilada. Con ello está dicho prácticamente lo más importante en cuanto a su significación didáctica. Pertenecen a esta clase de pruebas algunas como:

- «El pico más alto de Europa es...»
- «La distancia entre Madrid y Barcelona es de ... kms»
- «El primer satélite artificial se puso en órbita en el año ...»
- «La Constitución española se aprobó en Cortes en el año ... »
- « ¿Cuál es la región española con mayor riqueza en ganado vacuno? »
- « ¿Quién escribió La vida es sueño? »

Otro tipo de pruebas para la evaluación de conocimientos, son aquellas en las que se permite una mayor libertad de respuesta por parte del alumno, es decir, no se limita su contestación a una o dos palabras que ha de colocar en un determinado lugar, sino que se le permite una explicación razonada, aun cuando centrada en un término, en un asunto concreto. Por ejemplo:

- « ¿Qué son figuras cuadrangulares? »
- « ¿Qué significa el término emancipación? »

Dado que la variedad de contestaciones aun sobre el mismo asunto puede ser considerable, estas preguntas de tipo abierto exigen un mayor esfuerzo por el profesor para su calificación. Debe establecer mínimos y máximos: debe elaborar, antes de proceder a calificarlos, un esquema de posibilidades de contestación. Frente a esta dificultad presentan la ventaja de ser más fáciles de elaborar.

Las pruebas de elección presentan diversas modalidades. Una es la de dos alternativas, como por ejemplo:

- «La mayor altura del mundo se encuentra en el continente
- Asiático
- Africano »

De un tipo semejante son las pruebas llamadas de verdadero o falso, en las que ante un enunciado, hay que contestar con una de ambas alternativas:

- «Los ángulos de un triángulo suman 180° » V F

Tanto en unas como en otras, en todas aquellas pruebas de elección en las que se haya de determinar el sujeto por una de entre tan sólo dos posibilidades, existe el riesgo de que la respuesta no refleje el aprendizaje realizado más que en un 50%, ya que el otro 50% puede teóricamente deberse al azar. Por ello, resulta más conveniente la utilización de pruebas de elección múltiple, en las que a una determinada pregunta, acompañen varias alternativas de respuesta, siempre más de dos. Estas pruebas requieren mayor esfuerzo de elaboración, pues es necesario que los distractores, es decir, las alternativas que acompañan

a la autentica, pertenezcan al tema de que se trata y no sean excesivamente sugerentes de cierto o error, en cuyo caso ya no tendría efectividad la pregunta inicial. Por ejemplo:

1º «El río más largo de la vertiente atlántica es:
A El Miño/B El Dueto/C El Tajo/D El Guadiana»

Está claro que las alternativas realmente válidas son la B y la C, por lo incongruente de la A. Por lo que respecta a las B y C, nos podría quedar la duda de si pregunta se refiere sólo a España o a España y Portugal.

2º « ¿En qué regiones de España se producen más cereales?:
A. — Norte/B. — Centro/C. — Levante/D. — Sur»

En esta cuestión existe un distractor claramente no significativo, que es el A.

3º « ¿Cuáles de entre las siguientes se pueden considerar como causas de contaminación?
A. — Escape de coche/B. — Complejos industriales/C. — Centrales atómicas»

Nos encontramos ante alternativas válidas todas ellas. Se impone un cierto estudio del orden de contestación según la magnitud de estas alternativas.

Otra modalidad que pueden adoptar las pruebas de elección es la modalidad de emparejamiento. En ellas se presentan un determinado enunciado y dos relaciones de conceptos, entre los cuales existe correspondencia. El alumno ha de establecer dicha correspondencia enlazando los conceptos de una columna con los relacionados en la otra. Por ejemplo:

- «Unir mediante flechas los nombres de ciudades con los de los ríos que las atraviesan»

Lugo	Guadalquivir
Burgos	Ebro
Sevilla	Miño
Zaragoza	Alarzón

Pruebas para la evaluación de la capacidad de comprensión

En esta categoría cognoscitiva se trataría de evaluar la capacidad del individuo para traducir, interpretar o extrapolar datos a partir de una determinada situación. Al contrario que el caso anterior, aquí no interesan la repetición de las características y condiciones de la enseñanza. Se trata de sustituir al alumno en nuevos contextos, diferentes de aquel en el que aprendió, pero con unas características semejantes en cuanto a la dificultad del simbolismo utilizado y el contenido de aprendizaje.

Son utilizables por lo tanto, pruebas que requieran la traducción del lenguaje icónico a verbal o viceversa, de gráfico a verbal y al contrario, etc. O bien, situaciones en las que ante una pluralidad de datos, el alumno tenga que establecer distintos tipos de relaciones entre ellos a fin de obtener una adecuada interpretación de la situación que se proponga, o bien deducir consecuencias, posibles efectos futuros, etc.

Ejemplos de pruebas para la evaluación de la capacidad de traducción podrían ser:

1. «Ante el gráfico de producción de una determinada industria en los últimos diez años, explicar verbalmente las características de la evolución general de producción de dicha empresa en el periodo de tiempo mencionado»
2. Se proporciona al alumno datos sobre costes de mano de obra y materias primas, condiciones y funcionamiento técnico y administrativo, así como cuantos datos adicionales sean requeridos por él. Se le exige que:
 - a) Explique cuáles han sido las causas de la evolución empresarial.
 - b) Trace un proyecto prospectivo de la evolución que tal empresa seguirá en los próximos años y especialmente en cuanto a la producción.
3. Se proporciona a los alumnos datos sobre producción agrícola de una región en X número de años, debiendo ellos representar gráficamente dicha información numérica.
4. los alumnos han de elaborar gráficos climatológicos (sobre todo lluvia y temperatura) a partir de datos numéricos recogidos por ellos mismos a lo largo del año.
5. Tras traducir verbalmente los gráficos de una determinada región sobre climatología y producción agrícola durante determinado número de años, el alumno tratará de encontrar y establecer el mayor número de relaciones posibles entre los datos conocidos, obteniendo una visión prospectiva respecto a los próximos años.

Como es fácil de comprobar, mediante la primera prueba evaluamos la capacidad de traducción. En la 2.a se evalúa la capacidad de interpretación y en la 2.b la extrapolación. En la tercera, se vuelve a evaluar la capacidad traductora, esa vez pasando del lenguaje verbal a la representación gráfica contrariamente que en el primer ejercicio. El cuarto ejercicio constituye asimismo una prueba de la capacidad de traducir, si bien con la novedad de que se realiza sobre datos que los propios alumnos han recogido previamente. La quinta prueba, finalmente, trata de detectar la capacidad de interpretación del alumno a partir de la relación de datos que previamente haya traducido y también se trata de evaluar en esta prueba la capacidad de extrapolar datos a partir de los que se tienen.

En esta segunda categoría (de la comprensión) evaluamos conductas mentales de una envergadura muy superior a las del simple conocimiento. A la rememoración, que se evaluaba con anterioridad, ahora se superponen las actividades cognoscitivas superiores de la traducción, la interpretación y la extrapolación. Las pruebas son sustanciales diferentes, pues en ellas se parte no ya del recuerdo sino de la supuesta posición de datos para proceder a actuar a partir de ellos.

Pruebas para la evaluación De la capacidad de aplicación

¿Qué aspectos se van a evaluar en esta categoría de «aplicación»? B. Bloom enumera las siguientes conductas susceptibles de evaluar.

1. Determinar qué principios o generalizaciones son apropiados o pertinentes para afrontar un nuevo problema.
2. Reformular un problema a fin de determinar qué principios o generalizaciones son necesarias para su evaluación.
3. Especificar los límites dentro de los cuales un principio o generalización particular es verdadero o relevante.
4. Reconocer las excepciones a una generalización particular y las razones para ellas.
5. Explicar nuevos fenómenos en función de principios o generalizaciones conocidas.
6. Predecir lo que sucederá en una nueva situación mediante el uso de principios y generalizaciones apropiados
7. determinar o justificar un determinado curso de acción o una decisión en una nueva situación mediante el uso de principios y generalizaciones adecuados.
8. Indicar el razonamiento que emplea en apoyo de la utilización de uno o más principios ante un determinado problema¹

La característica principal de las pruebas de aplicación es la de verificarse en contextos diferentes de aquel en que tuvo lugar el aprendizaje. Se trata de comprobar la posibilidad de transferencia del aprendizaje. De alguna manera, toda la institución escolar, como ha afirmado algún autor, reposa en esta posibilidad de transferir, de aplicar en otros contextos (disciplinas, ambientes, vida) las principales generalizaciones asimiladas a través de la instrucción. En su obra *Evaluación del aprendizaje*. Boom y otros se extienden ampliamente en este tipo de pruebas, pienso que en razón de la importancia relativa de las mismas. Como ejemplos significativos podríamos citar algunas pruebas como:

1. «Explique Vd. Por qué razones didácticas las imágenes ocupan un importante espacio en los libros de texto de los primeros cursos»
2. «Realice Vd. Una programación de actividades para 6º curso de EGB durante una semana, teniendo en cuenta:
 - a) Todas las áreas de aprendizaje
 - b) La necesaria combinación de períodos de trabajo y descanso
 - c) El diferente índice pronogénico de cada actividad
 - d) Otras consideraciones didácticas que estime de interés
 - e) La adecuada distribución de actividades de carácter individual y socializado».
3. «Realice Vd. Una relación secuencial de la serie de principios matemáticos necesarios para la resolución del siguiente problema:
“Calculo de la superficie de un terreno de tal forma y dimensiones»
4. «Explique Vd. En que condiciones una lección magistral retransmitida por TV puede resultar excepcionalmente más eficaz que otra realizada directamente»
5. «Alumnos que han estado sometidos durante años a un sistema autoritario-rígido de educación se ven de pronto en un ambiente liberal ¿Cuáles serán sus reacciones?»
6. «En la programación que ha realizado Vd. en la prueba 2 ha diseñado una determinada metodología ¿Puede explicar las causas que le han inducido a proyectarla así y no de otra forma?»

Las seis preguntas anteriores son, como se puede comprobar, eminentemente abiertas. A través de ellas podemos detectar en qué manera el alumno es capaz de aplicar determinados principios a una situación problemática más o menos nueva para él. Por

¹ BOOM, B.: HASTINGS, J. y MANAUS, G.: Ob. cit., pag. 245

ejemplo, ocurre así en la prueba segunda donde el alumno, se supone de ciencias de la Educación, ha de poner en práctica numerosos principios sobre la programación, metodología e higiene en el trabajo.

En el primer ejemplo, sin embargo, no se trata de aplicar principios sino de descubrir cuáles están aplicados ya en una determinada situación, mejor, en este caso, en un material didáctico.

En el tercer ejemplo de prueba, no se trata de comprobar la habilidad del alumno para aplicar principios, sino más bien su capacidad para prever cuáles serán necesarios en un momento determinado para resolver un problema.

En el cuarto ejemplo se trata de evaluar la capacidad del alumno para hallar los límites de aplicación de principios determinados. Es evidente que la elección directa causa unos efectos pedagógicos superiores a la retransmitida. Pero este principio tiene sus limitaciones y no siempre se cumple. Pues bien, en este caso el alumno tendría que razonar sobre los motivos de excepción a dicho principio.

En el quinto ejemplo se prueba se trata de evaluar la capacidad del alumno para prever, en virtud de la aplicación de ciertos principios didácticos-psicológicos, los efectos de una situación didáctica actual de cara al futuro.

Y finalmente, en el sexto ejemplo se evalúa, no ya la capacidad de aplicación de principios, sino la capacidad del alumno para explicar las razones que le han movido a dicha aplicación.

Así pues, siguiendo a Bloom, podemos comprobar cómo en esta categoría de actuación cognoscitiva existe una amplia variedad de matices. A la simple aplicación de principios, generalidades, se ha de sumar la capacidad del alumno para llegar a identificar principios ya aplicados, límites de la actuación de los mismos, previsión de futuro en virtud de aplicación de principios e incluso explicación de causas que han determinado la elección de principios.

Evaluación de la capacidad de análisis

Constituye el análisis de una tarea de envergadura superior a las anteriores, concretamente a la comprensión. Una vez conocida la naturaleza de un determinado contenido, antes de poder proceder a realizar una evaluación del mismo, es necesario profundizar a sus elementos, relaciones de éstos y estructura. Esta actividad de profundización en los contenidos de aprendizaje es precisamente lo que se recoge con el título genérico de *análisis*. Dentro de ella, Bloom incluye conductas de una complejidad progresiva como:

- Identificación y clasificación de los elementos de un contenido (por ejemplo, de una comunicación, un texto literario, etc.) de acuerdo con diversos criterios previamente establecidos.
- Identificación y clasificación de elementos al igual que en el caso anterior, pero esta vez no directamente explícitos sino conocidos a través de una diversidad de indicios.
- Estudio de las relaciones que se puedan dar entre los diversos componentes de las comunicaciones o las unidades de estudio.

- Análisis de los principios de las estructuras o los propósitos particulares sobre los que se basa la organización del objeto de análisis.

Así pues, las pruebas para la evaluación de esta habilidad genérica de analizar, pueden entrarse en los anteriores aspectos particulares de la misma. Estas pruebas, según el mismo autor, han de atenerse a unas características fundamentales como:

- a) Realizarse en contextos, con materiales diferentes a aquellos en los que tuvo lugar el aprendizaje.
- b) Poner a disposición de los alumnos el material de análisis de tal modo que lo puedan manejar mientras dure la prueba.
- c) Contrastar los resultados del análisis realizado por el alumno con el que realice una serie de expertos, de personas competentes en el tema objeto de análisis.

Puede servir de ejemplos para la evaluación en esta área algunas pruebas como:

Se les proporciona a los alumnos un texto literario, científico, social... y se les pregunta:

- a) « ¿cuáles son los enunciados subjetivos y cuáles los objetivos?»
- b) « ¿Qué argumentos se pueden catalogar como fácticos y cuáles como normativos?»
- c) « ¿Cuáles son los supuestos no enunciados que contiene el texto?»
- d) « ¿Se puede distinguir entre argumentos principales y derivados o entre enunciados y conclusiones?»

La dificultad del análisis que deba realizar el alumno está en relación con la naturaleza del texto proporcionado según que en el mismo se encuentren claramente explícitos los argumentos o que deban inferirse a través de una pluralidad de indicios.

Una prueba centrada en la evaluación no ya de elementos sino de relaciones podría tener como punto de partida un documento similar al señalado anteriormente y contener cuestiones como:

- a) Reconocer los propósitos y presupuestos del autor
- b) Reconocer los supuestos que apoyan determinados juicios
- c) Identificar argumentos fundamentales y otros accesorios o secundarios.
- d) Determinar relaciones causales entre ciertos elementos, fenómenos.
- e) Reconocimiento de las posibles relaciones entre argumentos explícitos y otros omitidos

En la evaluación del análisis de principios y estructuras de organización, las cuestiones a proponer se centrarán en aspectos como:

- a) Determinar el propósito fundamental del autor al escribir dicha obra
- b) Conocer cuál era su estado de ánimo al hacerlo.
- c) Explicar las características de la forma de una obra literaria.
- d) Encontrar la estructura de organización que da coherencia artística a una obra de arte.

Las cuestiones que se plantean al alumno pueden ser de carácter abierto como las anteriores, o bien más estructuras como en el caso de ofrecer con cada pregunta una serie de alternativas de respuesta, entre las que poder elegir. A partir de un estímulo determinado, pueden plantearse cuestiones relativas a los tres tipos de análisis que hemos mencionado. Así, a partir de la siguiente obra de miguel de Unamuno, podríamos construir las siguientes pruebas:

CASTILLA

Tú me levantas, tierra de Castilla
en la rugosa palma de tu mano
al cieno que te enciende y te refresca,
al cielo, tu amo.

Tierra nervuda, enjuta, despejada,
madre de corazones y de brazos,
toma de presente en ti viejos colores
del noble antaño.

Con la pradera cóncava del cielo
lindan en torno tus desnudos campos
tienen en ti cuna el sol y en ti sepulcro
y en ti santuario.

En todo cima tu extensión redonda
y en ti me siento al cielo levantado
aire de cumbres es el que se respira
aquí en tus páramos.

¡Ara gigante, tierra castellana,
a ese tu aire soltaré mis cantos
si te son dignos bajarán al mundo
desde lo alto!

a) « ¿Cuáles se pueden considerar como enunciados subjetivos y objetivos?»

Podríamos contestar que con absoluta mayoría los subjetivos, si bien existen también enunciados reales objetivos como:

«Tierra nervuda, enjuta y despejada, madre de corazones y de brazos», y algún otro metaforizado, como «con la pradera cóncava del cielo lindan entorno tus desnudos campos»

Sería ésta una cuestión encaminada a evaluar la capacidad para el análisis de elementos, en este caso enunciados subjetivos y objetivos.

a) « ¿Enunciados fundamentales y accesorios?»

— A partir del propósito del autor, como enunciados fundamentales podríamos considerar:

- Castilla eleva al cielo (estrofas 1ª, 4ª, y 6ª).
- Castilla está unida al cielo (estrofas 1ª, 3ª, y 6ª).

— Como enunciados accesorios:

- Cielo amo de Castilla (1ª. Estrofa).
- Tierra nervuda, enjuta (2ª estrofa).
- Aire de cumbre el de los pájaros (5ª estrofa).

c) «¿Cuáles son las relaciones causales que se dan en la obra?»

Se da una relación esencial, presente en toda la obra entre el sentimiento de elevación al cielo, que el autor experimenta en esta tierra y la estrecha relación de la misma con el cielo

Finalmente, podríamos pensar en pruebas para la evaluación de principios y estructuras de organizaciones como:

d) «¿Cuál es el propósito fundamental del autor?»

—enlazar su tierra. Castilla, haciendo ver el maridaje de la misma con el cielo.

Las mismas preguntas podrían haberse formulado en un estilo más cerrado, ofreciendo una diversidad de alternativas de respuesta:

e) «¿cuáles de entre los siguientes, consideras enunciados objetivos?»

—Marca con una A los objetivos y con una B los subjetivos.

- Castilla me levanta.
- Su tierra es nervuda, enjuta.
- El cielo la enciende y la refresca.
- Castilla es madre de corazones y de brazos, etc.

f) «¿Cuáles de entre los siguientes son enunciados fundamentales» (márcalos con una A) y cuáles son accesorios? (márcalos con una B)

- Castilla está unida al cielo
- El cielo es el amo de Castilla
- Castilla es una tierra enjuta
- Castilla eleva al cielo.
- Es punto el aire de sus páramos, etc.

g) «¿Cuáles son a tu parecer las más patentes relaciones causales expresadas en la obra, de entre las siguientes?»

- Me levanta al cielo porque es tu amo
- Madre de corazones y de brazos —une el presente con el pasado
- La tierra rodea al cielo —Y es su cuna y sepulcro y sudario.
- Me levantas al cielo porque eres tierra elevada.
- Me levantas al cielo porque tu aire es limpio.
- Rezar en Castilla porque es altar gigante.
- Me levantas el espíritu al cielo, en virtud del simbolismo que representa la unión de la tierra con el cielo. Siendo ésta tan clara, tan diáfana, parece que en la tierra mi espíritu se eleva también con mayor facilidad.

h) «¿Cuál de entre los siguientes te parece como principal propósito del autor?»

1. Enlazar a Castilla.
2. Hacernos ver las características de la tierra castellana.
3. Destacar la unión del cielo y tierra en Castilla

4. Convertirla en altar para sus oraciones.

Pruebas para la elaboración de la capacidad de síntesis

Constituye la síntesis una de las categorías de aprendizaje más difíciles, elevadas, en la que el alumno, a partir de elementos, de materiales ya asimilados, llega a producir un nuevo resultado en forma de comunicación original, de un nuevo plan de actuación o de hallazgo de principios, de relaciones abstractas que expliquen la realidad de una forma más coherente. Nos encontramos, por lo tanto, ante una categoría que abarca actividades propiamente creativas del alumno en el plano de la expresión, de la acción y del pensamiento. Las pruebas destinadas a la evaluación de esta capacidad han de presentar al menos dos *características* fundamentales:

1. Realizarse en un contexto diferente de aquel en que tuvo lugar el aprendizaje.
2. Contar para su realización con todos los materiales de trabajo que puede necesitar y que libremente utilizará en la forma y con el tiempo que crea conveniente.

Constituye, pues, un tipo de prueba muy diferente al del resto de los exámenes tradicionales. Se trata de suprimir aquí todo factor tensional en cuanto a fuentes de información o de límites temporales. Los datos de partida no constituyen elementos de interés sino en la medida en que pueden servir para la elaboración de un nuevo producto. Por ello, no se trata de comprobar si posee o no el alumno dichos datos sino de asegurar (proporcionándole material) que los pueda manejar de cara a nuevas producciones. La situación en que se encuentra el alumno en la prueba de síntesis ha de ser lo más parecida a la situación normal de trabajo e investigación. Pruebas relativas a la elevación de la capacidad de síntesis podrían presentar características similares a las siguientes:

1. Expresar por escrito, de forma coherente, resumida y clara cuáles han sido los principales argumentos esgrimidos en un debate que sobre la institución escolar acaban de mantener seis profesores del centro durante hora y media y que han escuchado en su totalidad, elaborando al final una serie de conclusiones personales sobre dicha problemática. Para la realización de la prueba podrá contar con el propio debate grabado en cinta y cuantos informes adicionales solicite.
2. elaboración, por el alumno de un programa de instrucción contando para ello con una serie de fuentes de información relativa a:
 - Características de los alumnos a que se destina.
 - Niveles y áreas de contenido.
 - Materiales de instrucción.
 - Características ambientales, etc.
3. Elaboración de un anteproyecto de construcción de un puente, contando para ello con todos los datos necesarios en cuanto a situación, longitud y anchura, materiales, costes, etc.

4. el alumno elabora una hipótesis provisional, explicativa de las razones por las que se produjo el hundimiento del barco, contando para ello con los datos proporcionados por los medios de información durante las fechas inmediatas a las del suceso.

Pruebas relativas a la capacidad de evaluación

Se concibe la evaluación como el conjunto de actividades consistentes en la formulación de juicios de valor sobre determinados objetos, fenómenos, etc., en virtud de criterios internos al propio objeto de evaluación o bien en virtud de criterios externos. Entre los primeros se especifican la exactitud, precisión y cuidado en la elaboración del objeto, la consistencia de los razonamientos utilizados y su organización y los valores y puntos de vista que ha utilizado el autor. Entre los criterios externos se citan la comparación del objeto con otros o el establecimiento previo de normas generales y su utilización.

La actividad evaluadora ha sido situada por Bloom en la cima de su jerarquía de conductas, por cuanto que en ella se da una dificultad mucho mayor, requiriéndose también el previo dominio de las actividades correspondientes a los niveles inferiores de la taxonomía. En efecto, para poder emitir un juicio evaluador según los criterios que se han visto, es necesario que el sujeto sea capaz de analizar, de aplicar criterios... etc.

Como condiciones generales que deben reunir estas pruebas se señalan las siguientes:

- a) El objeto de evaluación, en la situación de prueba ha de ser diferente a los utilizados para el desarrollo de esta capacidad en le período de instrucción.
- b) Durante la realización de la prueba, el alumno ha de disponer de los materiales que evalúa, ha de contar con ellos en todo momento.
- c) Conveniencia de que en la prueba de evaluación se pongan en funcionamiento varios de los criterios anteriormente mencionados.
- d) Los resultados de la evaluación que se realiza se han de comparar con los que realicen determinados expertos en el terreno que es objeto de evaluación²

Las pruebas para evaluar esta capacidad podrían presentar características semejantes a las siguientes:

El estudiante acaba de observar cómo un compañero efectúa demostraciones sobre cómo se realizan diversas reacciones químicas. Al final de la prueba emitirá un juicio evaluativo sobre dichas demostraciones en virtud de la exactitud, precisión y cuidado con que las realizó y estableciendo cierto tipo de comparación con prácticas similares que otros compañeros vienen efectuando regularmente.

² *Ibid.*, págs. 310 y ss.

Como es fácil de comprobar, en estas pruebas se cumplen las condiciones exigibles para la clase, como el contar con el material (demostración que observa con todo cuidado), utilización de diversos criterios internos (precisión, cuidado, exactitud) y externos (comparación). En cuanto a la primera condición, es decir, que el objeto de evaluación no haya sido utilizado en el período de aprendizaje, es de suponer que en efecto, las reacciones sean diferentes a las realizadas con anterioridad. Finalmente, es igualmente realizable la última a condición de comparar los resultados de la evaluación realizada por el alumno con la de uno o varios expertos.

Contando con datos relativos a todos los aspectos materiales, funcionales y personales de una institución escolar, el alumno ha de evaluar la misma contando como criterios para ello:

- a) Las disposiciones oficiales existentes.
- b) La comparación de dicha institución en sus elementos y funciones con otras del extranjero.
- c) El cumplimiento de los objetivos que en instrucción se propone dicha institución escolar.
- d) La coordinación existente entre factores personales, administrativos y demás recursos de cara al máximo de eficacia de cada uno de los mismos.

Se juzga, mediante esta prueba de habilidad del alumno para utilizar tres tipos de criterios: otra institución escolar con la que establecer comparación; una serie de criterios preestablecidos, la congruencia de los resultados de la institución con sus metas previamente establecidas y la coherencia o relación armónica entre sus diversos componentes.

Una vez vistas, aun cuando de manera aproximativa, el conjunto de pruebas que según Gagné, y más extensamente según Bloom pueden utilizarse para la evaluación de los aprendizajes del ámbito intelectual, reflexionando globalmente sobre sus características podríamos hacer hincapié en una serie de *aspectos* como:

Si bien, y como se señaló en principio, debe existir congruencia entre objetivos y metodología por una parte, y la naturaleza de las pruebas por otra, esto no quiere decir que las situaciones de evaluación deban ser exactas a las de aprendizaje. Muy al contrario, y a excepción de la primera categoría del aprender, es decir, los conocimientos, en los que sí debe existir una aproximación mayor, en el resto de los aprendizajes, es necesario que la situación de evaluación sea distinta a aquellas en las que el alumno experimentó para aprender. Sobre todo, cuando se trata de evaluar capacidades como la aplicación, en la que se intenta precisar exactamente la habilidad para la transferencia a nuevos contextos de los aprendizajes que el alumno adquirió en un momento determinado.

Aun cuando hemos ido estudiando pruebas específicas para cada aprendizaje también específico, en la realidad los aprendizajes se confunden, se relacionan tan estrechamente que no resulta significativo evaluarlos con independencia unos de otros. Esta realidad del aprender da lugar a que las pruebas de evaluación vengan a constituir baterías o conjuntos de cuestiones relativas a cada uno de los aprendizajes en particular. Esto, que

ahora lo referimos al ámbito intelectual, se podría aplicar asimismo a la necesaria relación existente entre este ámbito y el afectivo y psicomotor y por ello, hablaremos también de la vinculación de cuestiones relativas a la evaluación en estos otros ámbitos con el primero.

Las pruebas de evaluación para todas las categorías estudiadas pueden presentar dos aspectos: puede considerárselas como abiertas o cerradas según que exijan una elaboración personal de la respuesta o que ofrezcan cierto número de alternativas para la contestación, entre las que ha de elegir el alumno. Hemos visto esta doble posibilidad más extensamente en la cuarta categoría, de análisis. Concretamente, Bloom ofrece ejemplo de cuestiones cerradas para todas las categorías de aprendizaje en el ámbito intelectual. No obstante, parece necesario resaltar la importancia de cara a la evaluación de las aportaciones personales, en orden a exigir un creciente esfuerzo de actividad mental creadora del diseño de pruebas de carácter abierto, en las que se va a exigir un mayor esfuerzo de actualización y relación de conocimientos así como de reflexión por parte del alumno.

Evaluación de aprendizajes del ámbito efectivo

El ámbito de la efectividad, aun cuando se le ha venido concediendo gran importancia teórica a lo largo del tiempo, en la práctica, sin embargo, se ha visto notablemente abandonado dentro de la actividad didáctica. No obstante, querámoslo ver o no, las actitudes, intereses, valores de nuestros alumnos desempeñan un importantísimo papel de cara al aprendizaje y al desarrollo personal. Aun cuando no haya sido objeto de evaluación o tratamiento metodológico, se van a desarrollar en un sentido u otro, y van a influir poderosamente sobre la actuación total del alumno y más concretamente sobre sus aprendizajes de carácter intelectual. Actitudes hacia la colaboración, hacia el trabajo de cada una de las áreas del saber, interés más o menos intenso por determinado tipo de actividad, formación de un esquema de valores respecto a los contenidos culturales...son factores tan importantes que cerrar los ojos ante ellos constituiría cuando menos una negligencia culpable por parte del educador. La evaluación de la afectividad cuenta con obstáculos de al menos la misma magnitud que su tratamiento metodológico-didáctico. Durante mucho tiempo se ha considerado no factible dicha evaluación en virtud de una serie de motivos como:

1. La determinación de una serie de criterios para la evaluación de la afectividad haría caer a los educadores en un cierto tipo de adoctrinamiento. De hecho, se estaría propiciando una determinada perspectiva para el desarrollo de la afectividad. La afectividad se ha considerado en nuestra cultura occidental como perteneciente a la intimidad de la persona y se ha desarrollado la creencia de que cualquier intento de educación en este ámbito iba a enmascarar una intención manipulativa.

2. Quizás por estos y otros motivos de base, no existen estudios científicos de la envergadura de los realizados en el ámbito intelectual-cognoscitivo. Los mismos términos del área afectiva están llenos de ambigüedad, significando una misma palabra contenidos distintos para diversas personas. El desconocimiento de la evolución de los procesos afectivos es enorme. Se da la creencia de que tardan mucho en producirse y que por lo tanto, su evaluación se escapa de la actuación del centro escolar. Recientemente se ha

descubierto sin embargo, que ciertos procesos de aprendizaje afectivo pueden desarrollarse con mayor o la misma rapidez que otros de tipo intelectual. Consiguientemente, apenas se han elaborado ni desarrollado pruebas para la evaluación de la afectividad. En todo caso, la evaluación de la afectividad se caracteriza por su sentido formativo y por individualización y carácter cualitativo de los juicios que en ella se formulen. Si el alumno piensa que de su respuesta puedes seguirse una determinada calificación, puede desvirtuar sus manifestaciones afectivas hacia una serie de patrones considerados como ideales en el centro o por el profesor. De aquí, la necesidad de una evaluación individualizada, de aquí la necesidad de evitar la obsesión de la prueba en el alumno. Entre la diversidad de *instrumentos* a utilizar en la evaluación de la efectividad, Bloom menciona los siguientes:

Entrevista

Se verifica con la presencia inmediata, directa del alumno frente al profesor. Presenta dos modalidades: estructurada y no estructurada. En la primera, la entrevista consta de una serie determinada de preguntas sobre aspectos concretos de un tema e incluso las preguntas pueden ser cerradas, de complementación o de elección entre alternativas previamente diseñadas. Ni entrevistador ni entrevistados tienen apenas libertad de acción, ya que deben moverse dentro de unos márgenes previstos.

En la entrevista no estructurada, sin embargo, no se limita el mínimo repreguntas. Se prevé el diálogo sobre determinadas cuestiones, de uno o varios temas, pero tanto el número como la formulación precisa de cada pregunta, se dejan para realizarlo sobre la marcha, a medida que está teniendo lugar la entrevista.

Este último tipo requiere un mayor entrenamiento en el profesor y un conocimiento profundo de los temas, una cierta capacidad de intuición y rapidez mental para captar pequeños matices significativos.

Tanto uno como otro tipo de entrevistas presentan la característica de permitir conocer al profesor a través del lenguaje gestual y el tono de voz, matices, intereses, actitudes, sentimientos que no sería posible conocer tan completamente sólo a través de la palabra.

La entrevista no estructurada permite profundizar más en aspectos personales y ofrece una pluralidad de ideas, sugerencias, para la construcción posterior de otras pruebas ya más cerradas. Frente a estas grandes ventajas, la entrevista estructurada ofrece la posibilidad de un más rápido manejo estadístico de datos.

Cuestionarios

En ellos, normalmente, se presenta a los alumnos un conjunto de preguntas que han de contestar sin la ayuda o presencia directa del profesor como en el caso de la entrevista. Las preguntas de los cuestionarios pueden ser abiertas o cerradas. En las *abiertas*, el alumno ha de realizar un esfuerzo imaginativo para la elaboración de la respuesta que podrá tener mayor o menor extensión. Constituyen un tipo de preguntas especialmente aptas para captar intereses, actitudes, sentimientos, creatividad por parte de los alumnos. Preguntas de este tipo podrían ser, por ejemplo:

— «Si tuviera medios me gustaría investigar sobre.....»,

- «De hacer un viaje muy largo iría a.....»,
- «El mejor regalo que me podría hacer un amigo sería.....»,
- «Me gusta jugar a Porque.....»,

Cuando estas preguntas se formulan en primera persona como en el caso de las anteriores, la contestación del sujeto tiende a ceñirse a la imagen que quiere que los demás tengan de él. Si queremos liberalizar más su expresión haciendo que se manifieste como realmente es, resulta más apropiado utilizar en la formulación de dichas preguntas la tercera persona, pues de esta manera proyecta en un «él» imaginario sus gestos, intereses, actitudes y así desaparece el temor a ser juzgado por los demás. Por ejemplo, en lugar de

- «Cuando me encomiendan un trabajo.....», escribiríamos:
- «Cuando a Luis le encomiendan un trabajo.....»
- «Si tuviera un laboratorio propio.....», se convertiría en:
- «si Juan tuviera un laboratorio para el.....»

Además del indicado valor de una más completa expresión personal, las pruebas de final abierto pueden servir de fuente de conocimiento de interesantes datos para la elaboración posterior de otras cuestiones estructuradas como ocurría en el caso de la entrevista.

Los cuestionarios para la evaluación de actitudes pueden también estar constituidos por preguntas *cerradas*, con determinado número de alternativas que vienen a representar otros tantos grados de intensidad en el interés de la persona hacia aquello que constituye el núcleo de la pregunta. Así por ejemplo:

1. «creo que debería concederse una atención a la agricultura»:
 - a) Pequeña (inferior a otros sectores de producción).
 - b) Normal (parecida a otros sectores de producción).
 - c) Grande (superior a otros sectores de producción).
 - d) Enorme (muy superior a otros sectores de producción).

Otra modalidad de cuestionarios de preguntas cerradas, la constituyen las *escalas de intereses personales o profesionales*, que abarcan una pluralidad de actividades, a cada una de las cuales el alumno ha de calificar según una determinada serie de posibilidades como, por ejemplo:

- A: interesa/B: no interesa/C: es indiferente.
- Escribir a máquina
 - Redactar informes
 - Contestar oficios.
 - Información.

En relación, se podría decir que estrecha con las preguntas de carácter cerrado, se encuentra la técnica de *diferencial semántico* debida a Osgood en la que, dado un determinado concepto, se enumeran series de objetivos relativos al mismo, ordenados según pares contrapuestos entre los extremos de cada uno de los cuales se han organizado cinco posibles grados de libertad en la determinación de la persona que se evalúa. Así, por ejemplo, se pide al sujeto que se determine respecto a la calificación de conceptos como

«El mar, es para Vd...»,

1. Hermoso.....Feo
2. Estimulante.....Intimidador
3. Fuente de riqueza.....Causa de pérdidas
4. Extraño.....Próximo
5. alegre.....Triste

Pruebas proyectivas

Las pruebas de carácter proyectivo tienen por finalidad permitir la libre expresión de las actitudes, intereses del sujeto. Dicha libertad de expresión no se produce cuando el sujeto aparece comprometido en primera persona en la ejecución de actividades o en el compromiso con determinadas posturas. En estos casos, lo que ofrece es más bien una imagen ficticia, la que quisiera que los demás se formaran de él. En cambio, si no aparece directamente comprometido sino que refiere sus sentimientos a una tercera persona, entonces sí puede realizarse una expresión propiamente libre, completa. Esto es lo que se pretende precisamente en las pruebas proyectivas, la proyección de la propia personalidad en una tercera persona a fin de facilitar la libertad de expresión. Una manera de construir pruebas proyectivas es, en efecto, cambiar la primera persona de las preguntas por la tercer. Por ejemplo, en lugar de: «Me gustaría leer...», enunciar: «a mucha gente le gustaría leer...».

Otra modalidad de pruebas proyectivas lo constituye las preguntas abiertas, a las que hay que contestar con una respuesta de elaboración, de carácter personal y creativo.

Otra modalidad la constituyen aquellas pruebas que toman como base de partida, no ya una pregunta verbal, sino un estímulo visual: dibujo de escenas que cada sujeto ha de describir e interpretar a su manera. Ha de investigarse quizás una historia, de cuyas características resultará fácil deducir intereses, actitudes, valores, ideales de su autor...

El Dr. Ridruejo³, es autor de un sistema de evaluación de actitudes sociales, que él denomina Ecoproyección, y que consta como estímulo de dos series de láminas: una representativa de caras con distintos gestos y de muñecos también gestualmente significativos de diversas actitudes (satisfacción, miedo, extrañeza, etc.); otra serie de láminas representa diversas situaciones de la vida real, social, profesional, etc. Los alumnos a quienes se trata de evaluar, tras la contemplación de ambas series han de asociar cada una de las láminas de la segunda (situaciones sociales diversas) con las láminas de la primera serie (cara y muñecos significativos). De alguna manera, al contemplar las láminas de diversas maneras, el alumno manifiesta una determinada actitud sobre las mismas a través de las láminas que empareja a aquéllas.

Observación

Multitud de aprendizajes que el alumno realiza en el ámbito escolar se escapan a las posibilidades de evaluación mediante pruebas de carácter formalizado o bien, en estas situaciones se desvirtúa su auténtica realidad, como ocurre en las actitudes, intereses

³ El Dr. Ridruejo es profesor y director del Departamento de «Ecoproyección» en la Facultad de filosofía y ciencias de la Educación de la Universidad autónoma de Madrid.

valores, es decir, con los aprendizajes del ámbito de la afectividad. Las actitudes que un alumno muestra hacia sus profesores, compañeros, asignaturas, etc., el grado de interés que los objetivos del curso o la unidad y las actividades programadas, así como los recursos, etc., despiertan en él la satisfacción que se puede sentir por asistir y permanecer en el centro, etc. son ámbitos del comportamiento que se escapan a la evaluación dentro de una situación formalizada, pues en la misma el alumno tiende a ofrecer más la imagen que quiere que tengan los demás de él que la que en realidad él tiene o desearía tener. Aun cuando las pruebas proyectivas han venido a paliar en parte esta problemática, no lo han logrado en toda su extensión. Existen otros procedimientos para obtener datos sobre la conducta actitudinal y afectiva en general, del alumno. Dichos procedimientos se basan todos ellos en la observación del comportamiento discente. Nadie mejor que el profesor, que convive a diario varias horas con los alumnos para practicar este método general de recogida de datos que es la observación.

Existe una serie de técnicas específicas que pueden aportar precisión a los datos que en una primera observación espontánea, informal, pudiera recoger el docente. Entre dichas *técnicas de observación*. Lafourcade cita: las listas de control y las escalas de calificación (numéricas, gráficas y descriptivas).

Las *listas de control* vienen a estar constituidas por series de numerosas actividades relacionadas con diversas actitudes ante las cuales el profesor, a través de la observación de sus alumnos, ha de verificar si las cumplen o no, si aparecen o no aparecen en la conducta de los mismos. Por ejemplo, en la evaluación de la actitud de trabajo en equipo, el profesor podría utilizar actividades como:

— Escucha anteriormente a sus compañeros...	Si	No
— Presenta sus trabajos a tiempo...	Si	No
— Reconoce méritos en el trabajo...	Si	No
— Presenta ayuda cuando se le pide...	Si	No

Las escalas de calificación constituyen pruebas más precisas y complejas, por cuanto que no solamente se pide la constatación de la presencia o ausencia de un determinado rasgo, sino que se pide, además, el grado en que tal rasgo se da. En las escalas numéricas, dicho grado se presenta mediante un número cuya amplitud varía notablemente. Por ejemplo:

«Calificar los siguientes rasgos, del 1 al 5, según la siguiente valoración»:

1. Muy poco/2 Poco/3. Normal/4. Bastante/5. Mucho.

«Me gusta conservar los libros»

Las escalas gráficas representan, mediante trazos, los escalones de que se compone aunque también suelen representarse números letras. Por ejemplo:

« ¿Colabora de buena gana en los trabajos de grupo?»

X	X	X	X	X
<hr/>				
Nunca/	Poco/	A veces/	Bastante/	Siempre

En las escalas descriptivas, a cada uno de los escalones o posibilidades de evaluación se acompaña una breve descripción (de ahí el nombre) de cada grado de conducta a evaluar. En ellas se gana precisión para la recogida de datos. Por ejemplo:

«Actitud para la colaboración en el trabajo en equipo»,

- a) Siempre dispuesto a prestar su ayuda desinteresada para el logro de objetivos vinculados a su labor docente. Superpone el éxito del trabajo colectivo al suyo personal
- b) Evidencia conformidad al integrar grupos de trabajo
- c) Cooperera, pero sin mayor esfuerzo y sin mostrar muchas voluntad.
- d) Prefiere trabajar solo.
- e) No presta ni permite ayuda. Evita todo trabajo en común.⁴

Existe como una normativa general a tener presente en la construcción de estas escalas de evaluación. Según Lafourcade, es necesario considerar, entre otros, los siguientes aspectos⁵:

1. Una previa delimitación de objetivos a evaluar mediante la escala. Aquí es necesario advertir que dichos objetivos pueden también pertenecer a otros ámbitos diferentes al de la afectividad; se pueden evaluar aprendizajes intelectuales o psicomotores también. Si se trata de intereses, actitudes, etc., será preciso especificar claramente cuáles van ser objeto de evaluación. Por ejemplo:

- Actitud del alumno hacia el aprendizaje de las Matemáticas.
- Actitud del alumno en torno a las normas de organización del centro.

2. realizar una selección de rasgos de comportamiento concernientes al aprendizaje que va a ser objeto de evaluación. Por ejemplo:

Si nosotros queremos evaluar la actitud de un alumno hacia la Literatura, tendremos que hacer una previa especificación de conductas observables que reflejan dicha actitud. En este caso, algunas de las conductas a observar podrían ser:

- Invierte mayor tiempo en la realización de ejercicios de esta asignatura.
- Participa más activamente en las clase de Licenciatura, preguntando y expresando su opinión sobre diversos temas.
- Dedicar considerable cantidad de tiempo libre a la lectura de obras de carácter literario.
- Los libros que recoge en la biblioteca del centro son, en su mayor parte, de literatura.

3. La escala ha de tener una extensión congruente con la cantidad de rasgos a evaluar. Su número de ítems estará en relación con el número de conductas significativas del aprendizaje que se evalúa.

4. Es conveniente que la expresión de los ítems o cuestiones de la escala sea lo más clara y precisa posible, a fin de evitar confusiones en el registro y clasificación de datos.

⁴ LAFOURCADE, P. D.: *Evaluación de los aprendizajes*. Cincel, Madrid, 1972, pág. 168.

⁵ *Ibid.*, págs. 169 y ss.

5. Parece recomendable la intervención de varios expertos, tanto en la selección de rasgos como en la redacción de los ítems y en la evaluación final de los datos recogidos.

6. La calificación de los datos puede hacerse de la manera simultánea a su recogida cuando a cada ítem acompaña una graduación de los mismos en categorías que pueden oscilar de tres a siete por término general. Así, por ejemplo:

a) « ¿Cuánto tiempo invierte en actividades libres de esta asignatura? »

A B C D E

Mucho menos/ Menos/ Igual/ Mas/ Mucho más

b) « ¿Cuál es su grado de intervención en las clases de literatura? »

A B C D E

Muy bajo/ Bajo/ Normal/ Alto/ Muy alto

c) « ¿Qué cantidad de tiempo invierte en la lectura de obras literarias? »

A B C D E

Muy poco/ Poco/ Normal/ alto Muy alto

Observación incidental

Independientemente de una observación realizada mediante técnicas como las que acabamos de ver, a veces puede resultar interesante el registro de hechos especialmente significativos de aprendizajes realizados o de problemas de conducta. A lo largo de las numerosas horas que el profesor pasa con sus alumnos, sobre todo en los primeros niveles de enseñanza, tienen lugar, ocasionalmente, comportamientos, conductas especialmente reveladoras, cuyo registro puede servir para la posterior explicación de los motivos o causas que mueven al alumno a obrar de determinadas maneras, así como para poder aplicar un determinado tratamiento de recuperación o bien para poder realizar una orientación personal y escolar fundamentada.

Las *dificultades* que determinan que el profesorado utilice tan poco frecuentemente este sistema de recogida de datos, se cifran en:

— El elevado número de alumnos.

— La previsión institucional porque se cumplan los programas que relegan a segundo plano este tipo de actividades docentes en las que, en principio, parece que «se pierde tiempo».

— La falta de experiencia del profesorado en la recogida de datos mediante la observación y la posible contaminación de los mismos por la impresión subjetiva del docente.

Para realización de este tipo de observación incidental es necesario tener en cuenta una serie de condiciones como las expuestas por Lafourcade en sus obras sobre evaluación de los aprendizajes⁶.

a) Las conductas objeto de anotación han de poseer una cierta entidad pedagógica. Para Flanagan, los sucesos a observar han de poseer:

«Suficiente nivel de complejidad y significación como para que permita realizar algún tipo de predicción y análisis sobre la persona que lo ejecuta. La complejidad supone algo más que la simple respuesta a un estímulo individualizado. La significación implica el hecho de poner de manifiesto alguna característica de interés para el observador. De aquí la posibilidad de predicción y reflexión que propicia»⁷

b) Dichas conductas han de registrarse con datos concernientes al contexto en que tienen lugar. Se recomienda una breve y precisa descripción de lo ocurrido, pero si dicha descripción aparece totalmente aislada, no sería posible su interpretación. Una conducta del alumno puede explicarse de diferentes maneras y relevamos distintos datos según la situación en que tiene lugar.

c) No es conveniente que transcurra tiempo entre el suceso y su anotación, pues se podría desvirtuar la conducta. Pero a veces, en ciertas conductas, en las que puede influir la subjetividad del profesor, tampoco debe ser inmediata la anotación.

d) Parece conveniente que a cada anotación acompañe una primera interpretación provisional de la misma por parte del profesor, que será revisable en el futuro en un estudio de conjunto de las observaciones realizadas durante cierto período de tiempo.

e) Facilita el trabajo de recogida, conservación y manejo de datos, la utilización de fichas en las que por escrito se consignen una serie de datos como los siguientes:

Alumno	Fecha	Hecho contextualizado	Interpretación provisional
Observador	Lugar		
Curso	Hora		

Evaluación del aprendizaje psicomotriz

Existen numerosos aprendizajes de carácter psicomotriz o con un fuerte componente de esta clase de conducta. Basta extender la vista a través de cualquier currículo perteneciente a diversos niveles de enseñanza para poder percibirlo. Así por ejemplo, la lectura ora, la escrita, el dibujo de mapas u otras imágenes, el manejo de un instrumento de laboratorio, la realización de juegos y ejercicios gimnásticos, el manejo de cualquier tipo de equipo para el trabajo, como por ejemplo, los instrumentos meteorológicos, los útiles para el dibujo lineal, los aparatos de física, la coordinación verbal y gestual en representaciones teatrales, etc. Dada la importancia real de estos aprendizajes, cada vez, se siente como más

⁶ Ibid., págs. 354 y ss.

⁷ RODRÍGUEZ DIEGUEZ, J. L.: *Didáctica General I: Objetivos y evaluación*. Cincel, Madrid, 1980, pág. 354

urgente la necesidad de su programación y desarrollo sistemático y por lo mismo, de la evaluación de los resultados que el alumno va consiguiendo en este terreno.

La evaluación de los aprendizajes psicomotores puede centrarse en los procesos o bien en los productos. Normalmente, estamos acostumbrados a un tipo de evaluación sumativa, final, centrada en los resultados. Pero, precisamente para mejorarlos y saber exactamente a qué se deben, es necesario centrarse más en los procesos, en el conjunto de operaciones que el alumno realiza para poder obtener determinado producto de aprendizaje. Más que en ningún otro tipo de conducta, en la psicomotriz, es posible precisamente captar de manera sensible cuáles son los pequeños pasos que el alumno da en el proceso de aprender. Y esta posibilidad de observación de actividades debe utilizarse didácticamente para perfeccionar los resultados en una perspectiva formativa de la evaluación y la enseñanza en general. Podría afirmarse, por lo tanto, que las pruebas para la evaluación de los aprendizajes psicomotores deben basarse fundamentalmente en la observación y análisis de procesos y logros. Así por ejemplo,

En la evaluación de la escritura, un aprendizaje con un indiscutible componente psicomotriz, se consideran fundamentalmente dos factores: la rapidez y la calidad de los trazos. La rapidez se evalúa contando el número de letras que el alumno escribe en un minuto, recurriendo para ello a distintos procedimientos, fundamentalmente, la copia de un párrafo. La evaluación de la calidad requiere el estudio analítico de la letra desde la perspectiva de una serie de aspectos que según Freeman serían:

—inclusión / alineación / trazo / forma de las letras / espaciado.

Según C. T. Gray es necesario considerar nueve aspectos en la evaluación de la escritura y conceder a los mismos una puntuación ponderada:

1. Espesor.....3	6. Separación de palabras.....11
2. Oblicuidad.....5	7. Separación de letras.....18
3 Tamaño.....7	8. Limpieza.....13
4. Alineación.....8	9. Formación de las letras.....26
5. Espaciamiento de líneas.....9	

Existe, asimismo, otro procedimiento para la evaluación de la escritura, basado en la comparación de la realizada por cada alumno con una escala o serie ordenada de muestras de escritura, previamente recogida como representativa de una multitud de protocolos y dotada de carácter estandarizado. Cuando no se cuente con instrumentos de esta clase, el mismo profesor puede proceder a elaborar una escala con las escrituras de sus alumnos. Este procedimiento, aun cuando no tenga posibilidades de universalización, puede constituir un valioso útil para la calificación a efectos internos. El procedimiento de evaluación consiste entonces en comparar una determinada escritura con la escala previamente confeccionada y otorgarle la calificación correspondiente al grado de la escala con el que guarde mayor semejanza.

Nos hemos referido a procedimientos para la evaluación del producto, es decir, de la escritura ya realizada. Ahora bien, es necesario hacer hincapié en la importancia de la evaluación de las actividades que el alumno realiza para «desembocar» en ese producto. Así, existen una serie de operaciones previas como:

- Postura del cuerpo.
- Forma de coger el útil de escritura.
- Colocación del papel.

- Sentido de trazado de las letras
- Distancia de la vista a la mesa de trabajo

Todas estas actividades, que forman parte del proceso de escribir, deben ser objeto de un tipo de evaluación formativa. El profesor observa atentamente, sobre todo en los comienzos del aprendizaje de la escritura, cómo las realiza cada uno de sus alumnos a fin de proporcionarle la orientación necesaria en cada caso. De esta manera está previniendo la calida de la escritura y evitando la aparición de defectos difíciles de subsanar *a posteriori*.

Un importante instrumento para la evaluación del aprendizaje psicomotriz son, por lo tanto, las listas de control, constituidas por series minuciosamente detalladas y ordenadas de las actividades que el alumno debe realizar en cada proceso de aprendizaje y también por series de características que deba reunir el producto o resultado. Así por ejemplo, según Lafourcade, en el proceso y resultado hay que considerar.

PROCESO (Encuadernación de un libro)

a) Desmontaje de los cuadernillos de un libro para encuadernar:

1. Corta con precisión el hilo que une los cuadernillos.
2. Desprende los cuadernillos sin dañarlos.
3. Limpia bien el lomo para quitarle restos de cola o tiras de papel
4. Recortar dos cuadernos suplementarios de cuatro páginas cada uno.
5. Coloca uno al principio y otro al final.

b) Preparación del encuadernado:

1. Prensa fuertemente los cuadernillos para que los lomos presenten una superficie plana
2. Traza sobre el lomo cierto número de cortes serruchados.
3. Cuida de que no sean muy profundos para que no se vean cuando se abre el libro.
4. Coloca el mazo de cuadernillos en el telar.

PRODUCTO (Verificar El acabado de una repisa)

1. Las uniones de las ensambladuras se adosan perfectamente.
2. La superficie está bien alisada.
3. Los bordes mantienen una línea perfectamente lograda.
4. Los ángulos han sido cortados con precisión
5. El encolado no afecta al aspecto general del mueble.
6. El pulimento para el lustre final es el más adecuado⁸

Muchos otros ejemplos de lista de control podríamos recordar. Por ejemplo, las operaciones cuidadosas que ha de realizar un atleta, previas a un salto de altura o al lanzamiento de martillo; las operaciones a realizar para colocar las manos y mover adecuadamente los dedos en la escritura a máquina; los movimientos a realizar por un músico para tocar un determinado instrumento musical, etc. Es posible pensar en la utilización de instrumentos técnicos para el grado y reproducción de secuencias como la

⁸ LAFOURCADE, P. D.: *Ob. cit.*, pág. 164

TV y la fotografía. La repetición a cámara lenta, la observación por el propio alumno, de su manera de actuar pueden convertirse en un importante medio de perfeccionamiento.

Es posible pensar, asimismo, en la programación de pruebas formales específicas, para poder apreciar el grado de habilidad psicomotriz de nuestros alumnos. Estas pruebas van a tener diversas características según la materia de aprendizaje, Lafourcade señala para una prueba de habilidad en el trabajo con madera, los siguientes componentes:

- Lista de operaciones necesarias.
- Lista de herramientas
- Lista de materiales.
- Instrucciones para los alumnos.
- Análisis del trabajo realizado mediante una hoja de calificación previamente elaborada.

El área pretecnológica, la expresión dinámica, son ámbitos de clara predominancia del componente psicomotriz y, por lo tanto, en ellos tiene justificación más que en ningún otro, la elaboración de este tipo de pruebas formales.

Concretamente en el terreno de la educación física y refiriéndonos al nivel elemental de enseñanza, un posible programa de evaluación formal podría presentar los siguientes elementos:

1. Prueba inicial en diversos tipos de ejercicios (carrera, saltos, trepa, gimnasia...), realizados a comienzos de curso para determinar el estado inicial del alumno.
2. Prueba final en los mismos capítulos para establecer una comparación y verificar si se ha producido algún tipo de progreso.
3. Establecimiento de una puntuación referida al grupo, mediante la conversión de las puntuaciones individuales en puntuaciones típicas.
4. comparación de los resultados individuales y de grupo con los niveles considerados como mínimos y apropiados para el nivel de que se trate.
5. en cada una de las pruebas, inicial y final, el profesor debe tomar datos a través de la observación, tanto del proceso o ejecución como de las marcas obtenidas que le permitan calificar a sus alumnos en una serie de aspectos de desarrollo psicomotor, como los siguientes:

—Agilidad/Flexibilidad/Resistencia/Coordinación/Equilibrio/Velocidad/Fuerza/Progreso general.

6. Para la realización de la prueba inicial, parece recomendable seguir una serie de pasos como:

- Explicación teórica de las características de la misma.
- ejecución práctica por el profesor o un monitor a la vista de los alumnos.
- Periodo de ensayos por parte de éstos.
- ejecución de la prueba para calificación.

En la prueba final, ya no será necesaria esta tarea demostrativa previa, pues habrán sido practicados los ejercicios a lo largo del curso.

Independientemente de estas áreas de aprendizaje, en las que se daba una clara predominancia del componente psicomotor, en todas o casi todas las restantes materias didácticas existe con mayor o menor intensidad una manifestación psicomotriz. Así por

ejemplo, no cabe duda que en el aprendizaje de las ciencias físico-químicas y naturales se da dicho componente. Parece conveniente que en el proceso de enseñanza y de evaluación se les conceda la importancia que en la realidad tienen los factores psicomotores. Por ello, independientemente de las pruebas de evaluación específicamente diseñadas para la conducta psicomotriz, sería necesario que en toda prueba evaluadora se considerara este aspecto de la conducta. Así llegamos a la necesidad de superación del rígido esquema: pruebas intelectuales, afectivas y psicomotoras para propugnar un nuevo tipo de prueba global, centrada en las características reales del aprendizaje y mediante la cual se puedan evaluar todos los aspectos del mismo. Consideramos que en cada tipo de aprendizaje predomina una determinada clase de conducta, pero en todas ellas está presente la personalidad global del alumno. Por esto, parece empobrecedor y poco significativo continuar hablando de pruebas de conocimiento, de aplicación, de actitudes o de psicomotricidad, aun cuando a efectos de estudios sean necesario hacerlo y así lo hayamos practicado en las páginas anteriores.

Sin embargo, resulta absolutamente imprescindible dejar constancia de la necesidad de una evaluación global que responda al conjunto de factores que intervienen en el complejo proceso del aprendizaje, en el que en mayor o menor medida está presente toda la persona humana.

Síntesis

La evaluación constituye un elemento del proceso didáctico, estrechamente vinculado con todos los demás como los objetivos, recursos, metodología... Por ello, la naturaleza de las pruebas utilizadas ha de ser congruente con la del resto de factores de la instrucción. Tomando como punto de partida el aprendizaje, hemos visto como se pueden construir pruebas adaptadas a la evaluación del aprendizaje intelectual, afectivo y psicomotor. Para ello, hemos seguido los estudios de Gagné, y más específicamente, los de Bloom. De acuerdo con la naturaleza de la taxonomía elaborada por este autor y con la mayor abundancia de trabajos en el ámbito intelectual, nos hemos extendido más en él, aunque también hemos tratado de hacer ver la importancia y posibilidad de la evaluación en los terrenos de la afectividad y la psicomotricidad.

Terminamos siendo la necesidad de proceder a la elaboración de pruebas globalizadas, que tomando como punto de partida el conjunto de características reales de todo aprendizaje, en el que se incluye la personalidad total del alumno, recojan en un todo armónico y proporcional la evaluación de todos los ámbitos del comportamiento.

Bibliografía

- BLOOM, B.; HASTINGS, J. y MANAUS, G.: Evaluación del aprendizaje. Troquel, Buenos Aires, 1976.
- GAGNE, R. M.: Principios básicos del aprendizaje para la instrucción. Diana, México, 1977.
- GAGNE, R. M. y BRIGGS, L. J.: La planificación de la enseñanza. Trillas, México, 1979.
- LAFOURCADE, P. D.: Evaluación de los aprendizajes. Cincel, Madrid, 1972
- LEWY, A.: Manual de evaluación formativa del currículo. Voluntad/UNESCO, Colombia, 1976.
- RODRÍGUEZ DIEGUEZ, J. L.: Didáctica general I: Objetivos y evaluación. Cincel, Madrid, 1980.
- TENBRINK, T.: Evaluación. Guía práctica para profesores. Narcea, Madrid, 1981
- WES. C.: Investigación evolutiva. Trillas, México, 1975.

JIMÉNEZ, Jiménez. Bonifacio. “La evaluación de programas para la innovación” En: Evaluación de programas, centros y profesores. Madrid, Ed. DOE. Pp. 237-298

7

La evaluación de programas Para la innovación.

7.1. Introducción

Estamos en una situación de cambio en la formación. El medio en el que estamos inmersos exige a cada uno de sus ciudadanos unos conocimientos, un saber proceder distintos a los de tiempos pasados. Las nuevas tecnologías han dado lugar a otras formas de trabajo y a cambios en la vida cotidiana. Los adelantos en la comunicación han optimizado la relación entre personas de diferentes lugares y distintas culturas y ha dado lugar a un más amplio campo de trabajo. La escala de valores sociales, laborales, individuales básicos está variando y los intereses generacionales son muy distintos.

Esta realidad debe ser considerada en el ámbito de la formación, en cualquier nivel y, consecuentemente, se han de planificar y proyectar cambios parciales y/o globales. La enseñanza ha de ser funcional, cercana a las necesidades del medio, sea del mundo laboral, sea del ámbito social. En la formación inicial y continua se ha de aprender en función de las necesidades actuales y futuras.

Esta realidad da lugar a que en los centros, en las aulas, deban efectuarse innovaciones parciales o genéricas en la organización, en la gestión, en la planificación, en el ámbito de la enseñanza-aprendizaje (contenidos, técnicas de aprendizaje, evaluación, etc.). Es preciso innovar en función de las necesidades.

En pocos años la evaluación ha tomado un lugar destacado en el ámbito educativo. Este giro está en función de su conceptualización. En un primer momento evaluación equivalía a mediar, a conocer el nivel de un determinado aspecto educativo restringido su uso en el aprendizaje del alumnado.

En la actualidad evaluar es detectar cómo es una realidad educativa-alumno, grupo clase, centro formativo o educativo, planificación curricular, sistema educativo, etc. — sea desde una perspectiva cualitativa, cuantitativa o mixta, con el fin de tomar decisiones. La proyección de la evaluación es mucho más amplia y se extiende en todo el campo educativo. La evaluación es el medio por el cual valoramos y conocemos una situación educativa bien en su proceso bien en un momento determinado, pudiendo conocer la efectividad de la actividad.

El interés por la evaluación en las últimas décadas se ha acrecentado. En el Estado español se habla de la calidad de la educación. La LOGSE le confiere gran importancia al

presentar un título sobre calidad educativa. Han surgido directrices y normas para evaluar a los centros desde la perspectiva externa e interna. El ministerio de Educación y cultura recientemente ha pedido evaluaciones de la situación actual de la educación en el Estado español. A nivel de centros de formación también emerge la necesidad de evaluar un programa que se realiza y son las mismas personas y grupos que desarrollan proyectos educativos quienes piden que se evalúe el programa para conocer su efectividad.

En este capítulo se aborda la evaluación de programas desde la perspectiva de la innovación, ya que todo programa educativo en el que se introduzca una innovación necesita ser evaluado. Desde este punto de vista se realiza un breve análisis del concepto de evaluación de programas para la innovación y sus funciones; se consideran, los aspectos que deben tenerse presentes al querer desarrollar una evaluación de programas para la innovación; y se expone un modelo actual y experimentado como ejemplo.

7.2. La evaluación de programas para la innovación

En la introducción se ha presentado la palabra *programa* referida a un contexto educativo. Se conceptualiza como “actividad dirigida a solucionar un determinado problema considerando como tal la carencia de algún aspecto, la necesidad de introducción de alguna modificación para la mejora o la realización de alguna innovación”. Por consiguiente un programa puede incidir en uno o en diversos ámbitos de la educación (organización, didáctica, orientación educativa, etc.) y la amplitud de su proyección puede ser de muy diversa magnitud. Algunos ejemplos de programas educativos de innovación que se desarrollan en el momento son:

- Los programas que se llevan a cabo en centros educativos en la enseñanza-aprendizaje de la utilización de estrategias innovadoras para la resolución de problemas y cuya finalidad es la del logro de una mayor efectividad en la resolución de problemas.
- El programa RIF —Red de Instituciones de formación— que se inició en la década de los noventa con la cooperación de más de 500 instituciones o centros de formación de formadores de la Unión Europea. La finalidad de esta red internacional es la de realizar cursos de formación de formadores, crear materiales, realizar intercambios, etc., a fin de introducir vivencialmente y conceptualmente la dimensión europea en el futuro los profesores y, posteriormente, que ellos sepan incluirlo en los currículos educativos.
- El programa de Reforma Educativa del Estado español cuyo fin último es el logro de mayor calidad en la enseñanza y aprendizaje de los niveles reglados educativos no universitarios en función del cambio social y tecnológico.

Todos los programas que implican una innovación precisan de la evaluación.

La evaluación de programas para la innovación (EPI) se define como “el conjunto de actividades que se realizan para detectar y valorar la utilidad del programa en el contexto donde se proyecta y tomar decisiones, si son necesaria, para optimizar el programa sea en su totalidad sea en aspectos específicos”

Esta definición considera los aspectos técnicos (“Conjunto de actividades que se realizan”), *pragmáticos* (“detectar y valorar la utilidad del programa de innovación en el

contexto, es decir, la realidad formativa, social, cultural) y de *toma de decisiones* (“tomar decisiones, si son necesarias, para optimizar el programa sea en su totalidad sea en aspectos específicos”).

Debe considerarse la existencia de otras definiciones con otros enfoques. Las diferencias radican esencialmente en la visión del autor en relación a la finalidad de la actividad evaluadora: descripción de una situación (Tyler, 1942); la toma de decisiones (promover el cambio, la innovación, etc.: CIPP de Stufflebeam, 1972; Alkin, 1969); *quién emprende la toma de decisiones*, es decir, si son personas o entidades externas a las que desarrollan el programa (Patton, 1984), personas internas al programa, o el mismo grupo que lo ejecuta (House, 1983); *La relación o no con el contexto*: evaluación del programa en relación a sí mismo (Tyler, 1942) o evaluación del programa considerando el contexto y aspectos externos (Parlett y Hamilton, 1977; Santos Guerra, 1993); *La mayor incidencia en la metodología de la evaluación* (Mateo, 1990) o bien *en el proceso de decisión* (Talmage, 1982).

En el Estado español se han realizado en los últimos años diversas evaluaciones de programas para la innovación. Como ejemplos pueden citarse programas de formación de profesorado (Villar, 1997), la evaluación del alumnado de enseñanza primaria por el MEC (1997 a y b), y más recientemente la evaluación en el nivel de educación secundaria, efectuada a lo largo del año 1997-1998 y publicada en 1998. Anteriores a ellas hallaremos en diferentes Autonomías evaluaciones de programas experimentales de la Reforma Educativa (Andalucía, Cataluña, etc.).

Analizando estas evaluaciones se detecta que los planteamientos son diferentes. Unos son más cuantitativos, otros más cualitativos; unos detectan la repercusión de los nuevos planes en el rendimiento del alumno, mientras que otros consideran la evaluación más global incluyendo además la labor del profesorado, la organización del centro, los costos, etc. son proyectos distintos con un mismo objeto: detectar en unos programas para la innovación los aspectos positivos para potenciarlos y negativos para buscar nuevas vías de mejora.

7.3. Funciones de la evaluación de programas para la innovación

Las funciones de la evaluación de un programa inciden bien en el propio proyecto, bien en ámbitos externos a él. Colás y Rebollo (1993) presentan dos funciones básicas de la evaluación de programas: una función externa y social (Gross y Humpreys, 1985) y otra interna, en relación al propio programa (Sanz Oro, 1990).

En función de esta última clasificación, las funciones propias de la evaluación de programas para la innovación son:

- Funciones externas:
 - a) Informar y ayudar a comprender el programa innovador a quienes lo solicitan (expertos de la educación, profesores, padres, etc.) y a los responsables del mismo. Vencer las resistencias.
 - b) Favorecer los cambios positivos que aporta la innovación.
 - c) Evidenciar los aspectos negativos en función de nuevos cambios.
 - d) Detectar las relaciones entre el sector educativo en el que se desarrolla la innovación y el sector socio-cultural y el del trabajo.

- e) Fundamentar la toma de decisiones (rentabilidad, efectividad, economía de esfuerzos, mayor interés y motivación, etc.).
- Funciones internas:
 - a) Conocer explícitamente las dimensiones que abarca el programa independientemente en su finalidad y desarrollo.
 - b) Ayudar a comprender el programa y a identificar sus posibilidades y limitaciones.
 - c) Comprobar si el programa da respuesta a aquellas necesidades por las que se planificó.
 - d) Proporcionar información rigurosa sobre su desarrollo, plantear posibles readaptaciones para su mejora y/o desarrollar nuevos servicios o readaptar y/o potenciar los ya existentes.
 - e) Aportar datos que permitan una reflexión crítica sobre una determinada parcela del programa o de su conjunto.
 - f) Dar resultados y presentar propuestas para la toma de decisiones.

Estas funciones pueden verse reflejadas en evaluaciones de programas para la innovación como las que se han realizado o están en curso en el Estado español —“Evaluación del nuevo sistema Educativo en el nivel de enseñanza secundaria”, “La introducción de las nuevas tecnología en los centros de enseñanza”, etc.—, o bien en estudios de evaluación de otros países donde están desarrollados programas para la innovación y cambio educativo (reformas del sistema) —Francia, el Reino Unido, etc.—, o en proyectos que realizan organismos internacionales como la IEA (International Association of the Evaluation of Educational Achievement), la IAEA (International Association for Educational Assessment), la OCDE, etc.

7.4. Planificación y desarrollo de la evaluación de un programa para la innovación

Al proyectar una evaluación de programa para la innovación se ha de tener presente: los aspectos básicos implícitos en la evaluación de programas para la innovación y la secuencia para planificar y desarrollar la evaluación de programa.

7.4. I. Aspectos básicos

Éstos se concretan en las siguientes afirmaciones:

- a) La evaluación es un proceso de carácter dinámico.
- b) La evaluación tiene un carácter contextualizado. Recordemos que todo programa para la innovación está condicionado por el medio; por consiguiente ha de ajustarse a las condiciones reales en las que se aplica el programa.
- c) La evaluación ha tener presente la resistencia del desarrollo del programa.
- d) La evaluación debe aplicar procedimientos científicos, lo cual implica estrategias de diseño, una rigurosa recogida de la información y de obtención de los resultados, un profundo análisis de los mismos (validez y fiabilidad).

- e) La evaluación de programas para la innovación siempre supone la emisión de un juicio o valoración.
- f) La evaluación conlleva la toma de decisiones y la proyección en un después.
- g) La evaluación tiene una función ideológica dependiendo de la conceptualización de evaluación de programa y de la finalidad de la misma.

7.4.2. Fases de realización

En la realización de una evaluación de programa para la innovación se consideran tres fases a seguir:

- El análisis de los determinantes del proyecto: estudio del contexto.
- La realización de la evaluación de programas para la innovación.
- La toma de decisiones y proyección en el programa.

A) Primera fase: el análisis de los determinantes del proyecto. Estudio del contexto

Esta primera fase tiene presente cuatro núcleos.

NÚCLEO 1

Responde al estudio de los aspectos causales y de contenido de la evaluación específica. Estos aspectos los relacionamos a continuación.

1. Necesidad de la evaluación: es preciso definir la necesidad específica de la evaluación del programa para la innovación y la finalidad de la misma. La concreción de las respuestas al *¿por qué?* y al *¿para qué?* ayudan a determinar los aspectos centrales de la evaluación así como los indicadores más relevantes.
2. Momento en el que debe realizarse la evaluación: en los programas para la innovación la evaluación ha de plantearse al mismo tiempo que se planifica el programa de innovación. Pero, *¿cuándo* debe efectuarse?

La evaluación puede realizarse de forma continua durante el desarrollo del programa de innovación para incidir, si es necesario, en aspectos del propio programa, o en momentos puntuales para conocer su incidencia. En el primer caso la evaluación es de proceso. En el segundo la evaluación es llamada de producto.

En todo programa de innovación se puede utilizar una u otra, o bien las dos. En programas que se proyectan en una amplia población se utiliza mayoritariamente la evaluación de producto, mientras que en programas de innovación cuya proyección es un pequeño núcleo de población se utiliza más la de proceso o ambas.

3. Contenido de la evaluación: es necesario precisar el contenido de la evaluación, es decir, debe considerarse el *qué* se evalúa del programa. No hay duda de que, en general, el demandante —equipo o persona que pide la

evaluación del programa— indica este contenido. Pero no es suficiente y debe definirse en qué ámbitos educativos incide, los aspectos a tratar, la relevancia de cada uno de ellos, etc., a fin de poder determinar las áreas de evaluación, los indicadores, las variables.

4. Quién pide la evaluación de programa: es importante considerar quién es el demandante. Ciertamente no es lo mismo evaluar un programa de innovación para un ente público, que para técnicos y profesionales del ámbito educativos y de la formación, o para un grupo de padres que deseen saber sobre la formación de sus hijos. Aunque la evaluación podrá incidir más en un aspecto u otro según los intereses del demandante.
5. Proyección espacial de la evaluación del programa: el campo físico de evaluación del programa, el *dónde*, puede ser muy amplio y diversificado (en este caso se trabaja con muestras) o muy reducido. Ejemplos del primer tipo son las evaluaciones realizadas en las distintas Autonomías del Estado español con competencias propias en educación, o las efectuadas por el INCE en España sobre la experimentación y puesta en marcha de la Reforma Educativa; también son ejemplo de estudios amplios las de los programas de innovación que se llevan a término en redes europeas de centros de formación. Ejemplo del segundo tipo, de campo reducido, es aquella que se planifica y proyecta para un centro determinado y que, aun siendo un estudio muy complejo, tiene una proyección espacial muy concreta: un centro, un grupo clase.
6. Las resistencias al programa de innovación: en la planificación de evaluación de programas para la innovación, un aspecto importante a considerar es el de las resistencias que pueden presentarse en el desarrollo del programa de innovación. Éstas deben preverse al planificar el programa y al plantear la evaluación, sea una evaluación de proceso, sea de producto, o bien mixta. Las resistencias podemos clasificarlas en:
 - Materiales: elementos materiales –edificio, recursos didácticos, etc.- que redundan en una “frontera” en el desarrollo de programa.
 - Humanas: un grupo de las personas que desarrollan el programa no acepta la totalidad del programa de innovación o algún aspecto del mismo.
7. La existencia de otras evaluaciones de programas de innovación semejantes: es positivo para el equipo evaluador el conocimiento de otros programas de innovación semejantes así como de su proceso de evaluación ya que puede aportar especificaciones, instrumentos, proyección de técnicas útiles.

NÚCLEO 2

Se refiere a aquellas personas que coordinan, desarrollan y participan en la evaluación y a quienes desarrollan el programa.

1. Realización técnica de la evaluación (planificación, dirección, etc.); la cuestión ¿quién realiza técnicamente la evaluación? Es la que corresponde a este apartado. La evaluación de un programa para la innovación debe efectuarla una persona o un equipo con formación en este campo, que sepa

planificar, coordinar y dirigir grupos de trabajo y que tenga dotes de comunicación –verbal y escrita-. Este experto o quipo puede ser externo al programa, interno, en él (trabajan en el programa), o también mixto.

2. Quiénes participan en la evaluación. Definición de su rol. Este punto corresponde a las preguntas: *¿quiénes participan? ¿Cómo?*

En la evaluación de un programa de innovación puede colaborar un gran número de personas y entidades. Se clasifican según el trabajo que realicen en:

- a) Personas externas o internas al programa que colaboran con la dirección del mismo, en estudios de campo específicos, ayudando a definir instrumentos, obteniendo datos (entrevistas, reuniones de análisis y discusión, análisis de contenidos), etc. Estas personas conjuntamente con la dirección conforman el equipo técnico.
- b) Personas y entidades internas al programa y que participan contestando a entrevistas y cuestionarios, asistiendo a reuniones, entregando datos del programa.
- c) Personas y entidades externas al programa pero, por su trabajo o por la finalidad que tiene la institución, pueden aportar datos desde una perspectiva externa (empresas, asociaciones, gremios, sindicatos fundaciones, técnicos, políticos, etc.).
- d) Personas y organizaciones que colaboran en los trabajos de estudio: centros informáticos (introducción de datos, obtención de parámetros estadísticos, procesado de los textos de los informes, etc.), transcripción de cintas magnetofónicas, filmaciones, etc.

NÚCLEO 3

Este núcleo se refiere al aspecto más técnico de la evaluación.

1. Tipo de evaluación; el tipo de evaluación responde al paradigma – cualitativo, cuantitativo o mixto- y a la tipología del modelo.
Los modelos más comunes de evaluación de programas para la innovación, independientemente de clasificaciones, son los siguientes:
 - a) Evaluación de necesidades. Esta evaluación tiene como objeto la identificación de situaciones deficitarias e insatisfactorias y que reclaman una solución.
 - b) Evaluación formativa –evaluación de procesos-. Su finalidad es detectar situaciones positivas y negativas en el programa para originar cambios, en el momento que se precise, para su optimización. Por consiguiente los resultados que se hallan en la evaluación del programa y su análisis dan nueva información a los que desarrollan el programa y les permiten introducir modificaciones que a su vez repercutirán en la evaluación que se va realizando.

- c) Evaluación sumativa. Es aquel tipo de evaluación que se efectúa en un momento determinado del desarrollo del programa. Puede realizarse durante la ejecución del programa, después de haberse desarrollado una fase o al final del mismo.
2. Quién va a tomar las decisiones sobre el programa: es importante conocer *quiénes toman las decisiones* a partir de los resultados de la evaluación. Las propuestas pueden ser dadas a conocer en una u otra vertiente, en una u otra perspectiva y siempre dentro de la objetividad y de la ética de la evaluación. En sentido amplio las decisiones pueden ser tomadas por:
 - El técnico o quipo que ha efectuado la evaluación.
 - La autoridad política, los técnicos de la Administración, etc.
 - Los receptores directos del programa (profesorados, la dirección del centro, la inspección, etc.).
 - De forma mixta: la autoridad política y el equipo técnico conjuntamente; los técnicos y los receptores directos del programa, etc.
3. El informe puede realizarse desde diferentes puntos de vista según el demandante de la evaluación. Será preciso considerar distintas tipologías de informes, el mayor o menor grado de tecnicismos, la mayor incidencia en unos u otros aspectos.

NÚCLEO 4

La valoración del contexto de la EPI. La evaluación del contexto del programa evaluar nos determina la necesidad, la posibilidad y definición de los elementos esenciales. Para llevar a término este tipo de evaluaciones debe considerarse:

- La claridad y precisión de respuesta positiva hacia la evaluación de cada uno de los aspectos mencionados en los diferentes núcleos del contexto.
- La necesidad de la evaluación en sí misma.
- La posibilidad y las limitaciones (personales, económicas, materiales, etc.) de la evaluación del programa de innovación que se quiere evaluar.

Aceptada la evaluación del contexto, debe procederse a la segunda fase de la evaluación

B) Segunda fase: la realización de la evaluación de programas para la innovación

En función de lo indicado en la fase anterior –determinantes del contexto- el evaluador debe considerar la planificación y la realización de la evaluación del programa de innovación. Tres grandes apartados son los que la determinan.

LA PLANIFICACIÓN

1. Diseño y planificación. Modelo: al planificar la evaluación de un programa de innovación podemos acogernos a un modelo o definir un proyecto propio. En uno u otro caso es preciso delimitar los sectores y ámbitos de trabajo, las áreas a evaluar, determinar los indicadores y concretar, si es necesario -depende del modelo y tipología de evaluación y de las técnicas a utilizar-, los ítems. En el modelo se indica secuencialmente, y mostrando su dependencia, los objetivos específicos, la búsqueda o definición de los distintos instrumentos, la especificación de las actividades para la recogida de datos, la muestra o muestra, la formación, si la

hubiera, de los encuestadores, de los observadores, al análisis de estadísticos o de los datos obtenidos, la realización del informe y de las propuestas finales.

A través del modelo podrá detectarse el tipo de evaluación que se realiza, si es de proceso, de producto, o mixta, y los distintos factores intervinientes. La temporización también deberá estar reflejada.

2. La organización del equipo evaluador: la organización del equipo evaluador puede ser más o menos compleja dependiendo del programa para la innovación y de la amplitud del proyecto.

En la evaluación de grandes programas, la organización y gestión del equipo evaluador es de suma importancia, así como los nexos de comunicación que se establecerán. En evaluaciones de programas muy amplias, siempre es preciso considerar:

- Un equipo coordinador de todas las actividades: recoge los materiales, los instrumentos que se utilizan, establece relaciones entre las personas que desarrollan las técnicas de evaluación y que recogen los datos, entre los equipos que se ocupan de diferentes áreas en reuniones, etc.; se ocupa de aclarar aspectos, buscar elementos que se precisan, redactar el informe final.
 - Un equipo o varios de trabajo. Estos equipos pueden ser:
 - a) Especializados, es decir, definen y crean instrumentos, o realizan parte del estudio evaluativo, o se ocupan de la información de los datos y obtención de los estadísticos, etc.
 - b) Se ocupan de llevar a cabo la evaluación del programa de innovación en un espacio geográfico definido, efectuando gran número de actividades, al igual que los otros equipos.
3. La viabilidad: al efectuar la planificación de la evaluación de un programa siempre debe realizarse una valoración de dicha evaluación desde la perspectiva de la viabilidad de la misma (recursos humanos, materiales y económicos).

LA PROYECCIÓN DE LA PLANIFICACIÓN

Ejecución de la evaluación del programa de innovación.

1. Desarrollo-ejecución: durante el estudio de evaluación debe seguirse la planificación y programación realizada. No obstante, y aun cuando el modelo y la planificación pueden preverlo, ésta debe ser flexible y sobre todo en aquellos aspectos del estudio en los que las variables son dependientes.
2. La obtención de datos y análisis: los datos pueden ser de diversa índole y es necesario especificarlo tanto para efectuar su recogida, como para ejecutar los estudios descriptivos y estadísticos posteriores. Básicamente se tendrán presentes las clasificaciones de datos cualitativos y cuantitativos, y las de datos documentales (el proyecto curricular del centro, las actas de las reuniones, etc.), materiales (del edificio, de los recursos didácticos, biblioteca, etc.) y personales (del alumnado, del profesorado, de asesores, etc.).

El estudio posterior, si la evaluación acoge un amplio espectro de campos, será preciso hacerlo por ámbitos, áreas y sectores; en un segundo momento

se relacionarán los resultados de los distintos y se extraerán las conclusiones parciales; en la fase final se obtienen los valores más generales y globales.

LAS CONCLUSIONES Y PROPUESTAS PARA EL PROGRAMA DE INNOVACIÓN.

La evaluación acaba con las conclusiones y propuestas para la toma de decisiones. Las conclusiones deben dar respuesta a los objetivos generales y específicos planteados. Su exposición debe abarcar, en una primera parte, una visión global sobre el programa para la innovación y, luego, los aspectos específicos de sectores, áreas y ámbitos.

Las propuestas de cambio para una mayor efectividad y calidad del programa pueden introducirse conjuntamente con las conclusiones, ya que se derivan de ellas, o bien al final del informe.

C) Tercera fase: la toma de decisiones y proyección en el programa

Esta última fase corresponde a la toma de decisiones sobre el programa. Esta responsabilidad puede ser del equipo que realiza la evaluación del programa para la innovación, o de quienes desarrollan el programa o de órganos o equipos externos (autoridades políticas, administrativas, etc.). Muchas de las evaluaciones de programas para la innovación de tipo social que desarrollan las decisiones son tomadas por los políticos.

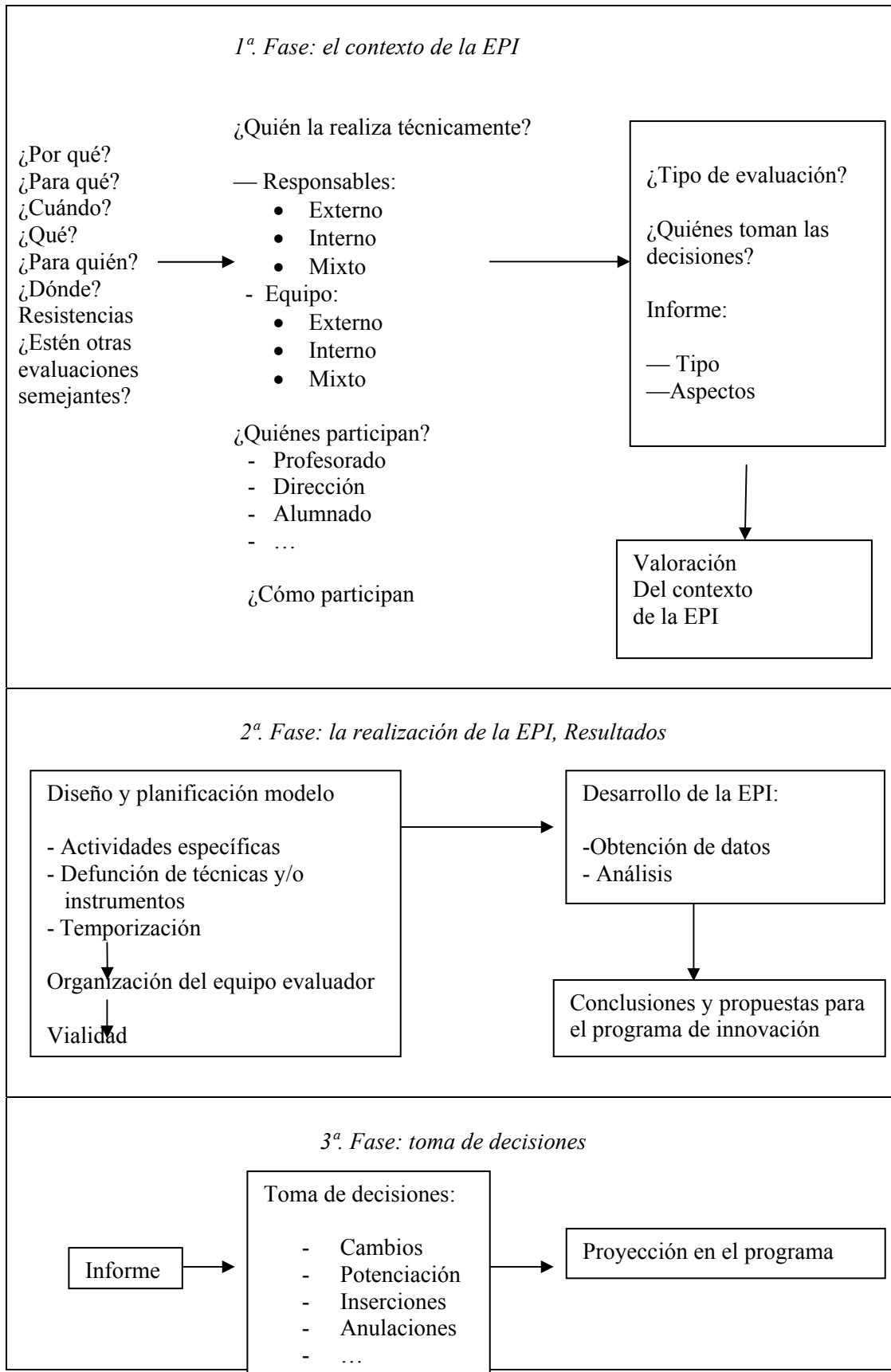
La figura 7.1 presenta el esquema de los aspectos que deben ser considerados en una evaluación de programas de innovación.

7.5 La evaluación de programas para la innovación en educación: el modelo EI

Para realizar una evaluación de programas para la innovación pueden utilizarse diferentes modelos creados por expertos o efectuar una adaptación al programa para la innovación que se desarrolla.

Los modelos de evaluación apropiados para programas de innovación podemos considerarlos dentro de la tipología de los de “toma de decisiones” (el CIPP de Stufflebeam, el CSE de Alkin, etc.), o entre los orientados a la “toma de decisiones por opinión” (el modelo de apoyo de Stake, el de expertos o críticos de Eisner, etc.), o los modelos de “oposición o adversos” (modelo de Owens y Wolf, etc.); también los modelos “pluralistas-intuitivos” (Stake, Parlett y Hamilton, MacDonald, etc.). Todos ellos se utilizan para constatar determinados aspectos de un programa educativo, para hallar la causalidad de determinados problemas, para potenciar aspectos positivos, etc.

El modelo el (evaluar para innovar) creado, desarrollado y experimentado por Bordas y Montaner, ha sido diseñado específicamente para evaluar programas para la innovación educativa. Se ha utilizado para valorar los programas experimentales de la Reforma Educativa que se han realizado desde 1984 hasta 1990 en Cataluña en el nivel de enseñanza secundaria (bordas, 1991; Rubies, Bordas y Montané, 1991). Su objetivo



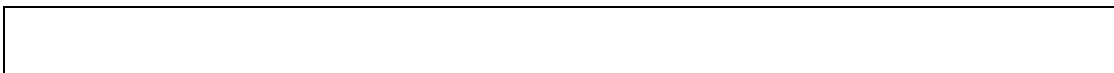


Figura 7.1. Fases en la evaluación de un programa de innovación

es evaluar programas para la innovación desde la perspectiva interna, de contextos y de continuidad a fin de mejorar la calidad educativa. Este modelo tiene presente:

- La comprensión y valoración del desarrollo y de los resultados de programas de innovación educativa.
- Las interrelaciones internas del programa y la incidencia en el contexto donde se desarrolla.
- La implementación futura de los ámbitos que representa el programa.
- La información para aquellos que deben tomar las decisiones para la mejora (directores, coordinadores, profesores, etc.).
- Un evaluador o equipo de evaluación (externo o mixto).
- Una participación activa en la evaluación del programa de las personas que trabajan en él y de aquellos a quienes la innovación educativa incide directa o indirectamente (empresa, campos sociales, etc.) aportando datos, documentos, indicando su visión, etc.
- Metodología y técnicas propias del paradigma cualitativo del cuantitativo (Cook y Reichardt, 1986). Ambas son consideradas como complementarias.

El modelo implica tres *sectores* de estudio. El *académico*, el *social* y el *profesional*. En cada uno de ellos se realizan análisis sobre la incidencia de la innovación educativa, obteniendo resultados que, relacionados con otros del mismo sector o de otros, dan lugar a las conclusiones y propuestas.

El *sector académico* tiene como finalidad conocer la proyección y efectividad del cambio en la realidad escolar y detectar, además, si el programa está dentro de las líneas actuales de la educación y la didáctica. Las áreas de estudio de este sector se deducen de los objetivos. A modo de ejemplo podemos indicar como grandes áreas la de “currículo”, la de “organización y gestión del centro”, la de “profesorado”, la de “orientación escolar y profesional y la tutoría”, la de “costos” de la enseñanza. Pero también pueden ser áreas de menor extensión como la de “desarrollo de la metodología para la mejora de la enseñanza-aprendizaje de una materia”, la “comunicación interna y externa del centro”, etc.

Los datos para el estudio de evaluación (fuentes) se obtienen de los alumnos, profesores, familiares de los alumnos, directores, orientadores, inspectores, administración expertos especializados en diferentes temas, etc., por medio de cuestionarios, entrevistas, observaciones, informes y dictámenes técnicos. Además es preciso hacer, en muchos casos, análisis de contenido de documentos que sean normativos, de formación del profesorado, de los alumnos, etc., que se han utilizado y/o se utilizan en el programa de cambio.

Cada una de las áreas de estudio tiene su propio diseño de trabajo diferenciado aunque está interrelacionado con los demás del sector.

El *sector social* da información valorativa del programa para la innovación que se desarrolla, desde diversos organismos y entidades sociales que conocen y tienen alguna relación con la educación y más específicamente con el aspecto de cambio que nos ocupa. En esta perspectiva pueden colaborar en la evaluación asociaciones de profesores, colegios

profesionales oficiales, ayuntamientos, asociaciones de vecinos, etc. las técnicas de recogida de datos más comunes son el cuestionario, la entrevista, los informes, la observación y las discusiones por oposición.

El *sector profesional* recoge información del mundo empresarial y del trabajo. La finalidad es conocer, por una parte, las necesidades de formación básica laboral y, por otra, su visión del programa que se evalúa. Colaboran en este sector las empresas de diferente magnitud (grandes, medianas y pequeñas), las cámaras de comercio, industria y navegación, asociaciones empresariales, etc. La presencia de este sector no siempre es necesaria. Es un ejemplo de este caso de estudio proyectado en la educación infantil.

La figura 7.2 muestra el esquema de este modelo de evaluación de programa de innovación.

Desde *la perspectiva organizativa* el modelo considera tres núcleos de trabajo:

- El núcleo *A* o *de estudio*, que tiene por objeto evaluar el planteamiento, desarrollo y proyección del cambio en la realidad. Este núcleo está compuesto por tanto grupos de trabajo como aspectos específicos de evaluación se consideren (currículo, organización y gestión, profesionalidad del profesorado, etc.) Cada grupo está coordinado por un director-coordinador experto en el tema. Los diferentes grupos de estudio se coordinan a través de los directores respectivos.
- El *núcleo B*, o núcleo consultor-asesor, lo componen grupos de trabajo, expertos y entidades que informan, asesoran y realizan dictámenes parciales sobre temas que se evalúan. Este núcleo lo pueden constituir profesores de centros donde se realiza el cambio, profesores de centros donde no se experimenta el cambio (con conocimientos sobre el tema que se evalúa), expertos, representantes de centros de investigación, representantes del ámbito empresarial, etc.
- El *núcleo C* o *coordinador* es el que dirige, coordina, facilita los datos o los medios para ser obtenidos, a los grupos y a las personas que los piden. Es el que redactará, a partir de todos los informes parciales, el informe final del estudio e indica las propuestas para la toma de decisiones.

La figura 7.3 presenta gráficamente la organización del equipo evaluador. El informe final considera los apartados siguientes:

- a) Conclusiones y propuestas para la globalidad del programa (el cambio que se desea realizar tiene un sentido global).
- b) Los aspectos específicos de las áreas de estudio que quedan determinadas en el núcleo A aun cuando se obtenga información de cada una de ellas en los otros núcleos.

Al mismo tiempo el informe debe dar una visión descriptiva del desarrollo del programa y de proyección futura (posibilidades, necesidad, relación con el medio, etc.). Es importante potenciar los aspectos positivos y representar propuestas de reforma para aquellos aspectos negativos, o menor positivos hallados.

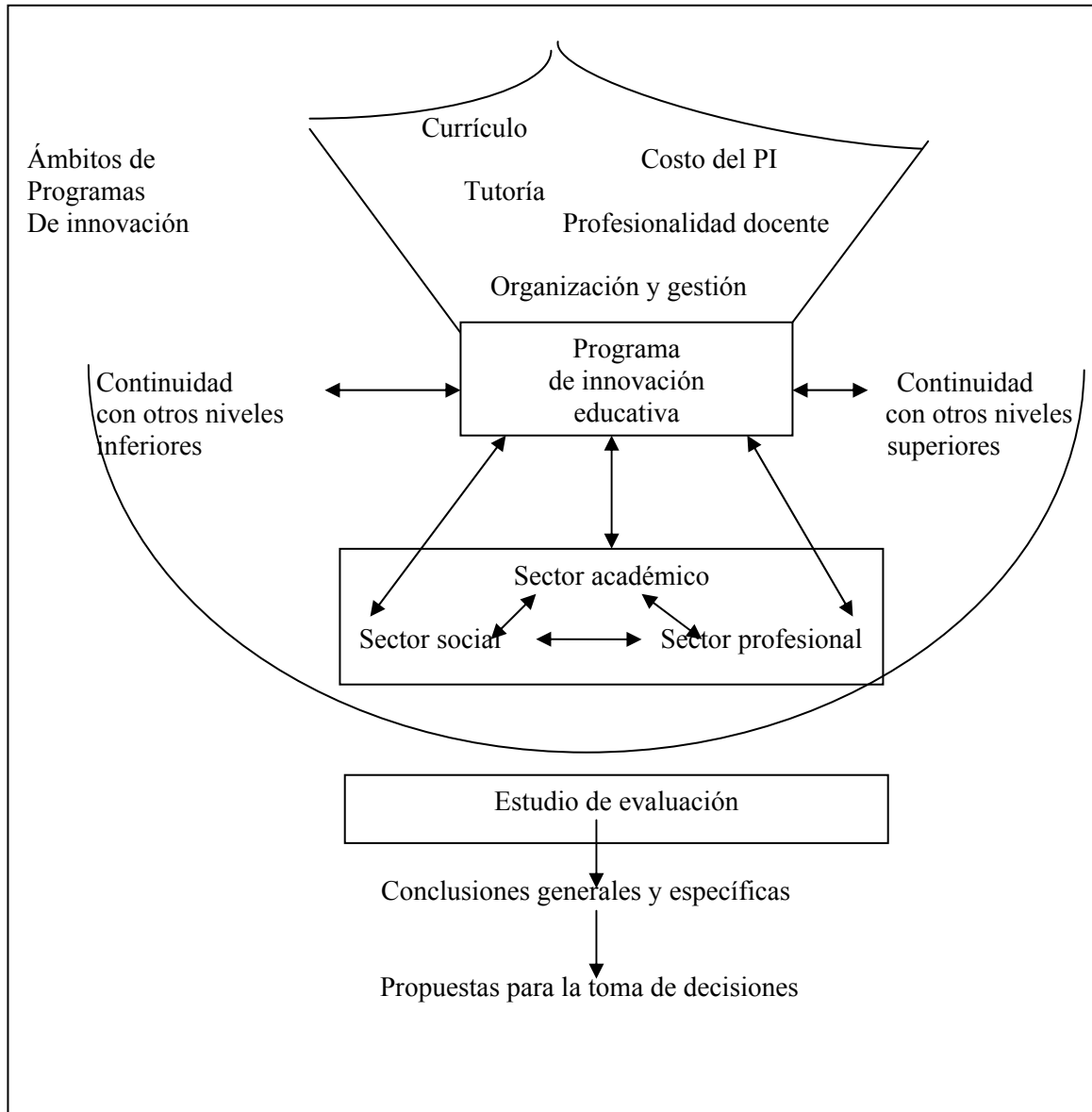


Figura 7.2 Esquema del modelo EI.

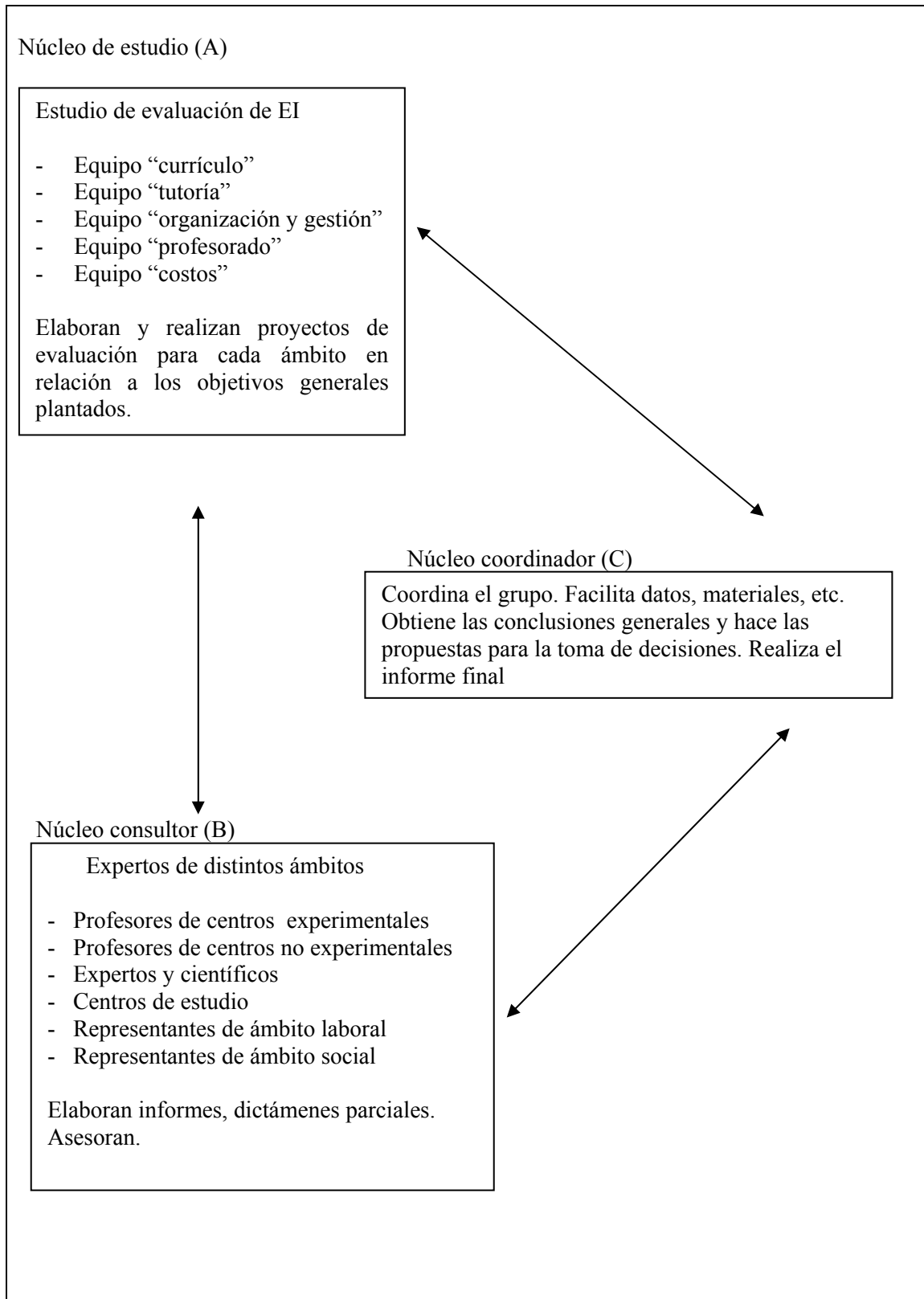


Figura 7.3. Esquema de la organización del equipo evaluador en el modelo EI

7.6 el modelo EI: proyección en una realidad

El modelo EI fu proyectado en la evaluación de los planes experimentales de la Reforma Educativa de la Enseñanza Secundaria en Cataluña. Los aspectos básicos del estudio se definieron de la siguiente manera:

- a) Finalidad y objetivos:
 - Detectar el cumplimiento de las directrices básicas y de los objetivos inicialmente planteados para el programa de reforma así como de los objetivos específicos para los planes experimentales de los diferentes niveles educativos que, desde el curso 1984-1985, se pusieron en funcionamiento de una manera progresiva.
 - Valorar en qué medida las propuestas de los panes se experimentan dan respuesta a las necesidades de la sociedad actual.
 - Considerar qué elementos de los planes experimentales analizados se pueden aplicar a todo el sistema educativo.

- b) Contenidos: los contenidos de evaluación quedan referenciados en cinco grandes áreas: la curricular, la de orientación escolar y profesional, la de organización y gestión de centros, la de las formación y profesionalidad de los docentes y la de costos económicos del nuevo sistema educativo.

- c) El equipo evaluador: el equipo evaluador es una comisión técnica externa de 39 miembros que desarrollan trabajos bien como expertos de algún ámbito, bien como representantes de un sector. En este equipo hay profesores de diversas universidades de centros de investigación, profesores de centros experimentadores y no experimentadores, representantes de la Administración y del sector profesional no docente (cámaras de comercio, de industria y navegación, agrupaciones de industrias —PIMEC, ASCAMM— y empresas) y del sector social (colegio profesional, centros de estudio). Cada uno de estos representantes colabora en la evaluación de diversas maneras:
 - Como directivo de un grupo de trabajo en un ámbito (núcleo A o *de estudio*).
 - Como asesor de los núcleos de trabajo y/o elaborando informes, aportando datos (núcleo B o *consultor-asesor*).
 - Como coordinador, haciendo trabajos de gestión, obteniendo datos de la Administración para el equipo, redactando el informe (núcleo C).

Algunos de los evaluadores trabajan en más de un núcleo.

- d) Participantes en la evaluación del programa: son participantes de la evaluación todas aquellas personas que trabajan en el ámbito de la Reforma Educativa y quienes directamente o indirectamente tienen alguna relación con el proyecto de cambio educativo. Por consiguiente profesores, directores, coordinadores, tutores, jefes de departamento, alumnos, inspectores, padres, personal de administración y servicios, representantes de ayuntamiento, de empresa, etc.,

participan en la evaluación aportando datos a través de cuestionarios, entrevistas, reuniones de discusión y juicio, etc.

- e) El tipo de evaluación, se define como externa y sumativa. Es de carácter positivo.
- f) Las técnicas: prioritariamente las técnicas utilizadas para la obtención de datos son: cuestionarios, entrevistas, reuniones grupales –descriptivas, de análisis y de juicio u oposición-, informes monográficos –científicos y técnicos-, análisis de contenidos de documentos, entrevistas, observaciones.
- g) Las conclusiones: las conclusiones –generales y específicas- para la toma de decisiones se obtienen directamente de los resultados de cada sector de estudio, a partir de relaciones entre los datos y los estadísticos de un mismo núcleo o por relación de datos y estadísticos de diferentes ámbitos.
- h) Los evaluadores, además de las conclusiones extraídas del estudio, presentan nuevas propuestas para la toma de decisiones.

Se valora el programa para la innovación desde cinco perspectivas o ámbitos que anteriormente se han indicado.

7.6.I. El ámbito curricular

Este ámbito o área de estudio desarrolla su propio esquema de trabajo. Los núcleos de estudio surgen de la normativa y de la concepción de la Reforma Educativa. Respecto a los tres sectores, podemos resumirlos en:

- El desarrollo del currículo: los objetivos generales de las etapas y de la continuidad entre etapas y entre niveles en los diseños curriculares.
- El currículo de ocho materias escogidas en función de su importancia, en cuanto a su novedad en el currículo y a su instrumentalización básica: estudio de los diseños curriculares, de su desarrollo y de su implementación en el aula.
- La individualización del currículo: optatividad, igualdad de oportunidades, individualidad en el aula.
- La interdisciplinariedad.
- La proyección nacional y europea en los diseños curriculares.
- La relación entre el currículo académico, las enseñanzas y los que se pide en el mundo del trabajo a los jóvenes de dieciséis y dieciocho o diecinueve años en distintos ámbitos laborales, fundamentalmente en “valores” y “habilidad y procedimientos”.
- La estancia de los alumnos en las empresas. Conexión del mundo académico con el mundo del trabajo.
- La preparación del alumno para iniciar sus estudios universitarios. Las pruebas de acceso.
- El abandono escolar en la nueva estructura organizativa y en función del nuevo marco curricular.

La figura 7.4 presenta las diferentes unidades de estudio y sus relaciones.

Cada uno de los trabajos se realiza a partir de un análisis de los planteamientos iniciales, de la puesta en marcha de los mismos y de los cambios efectuados a través de

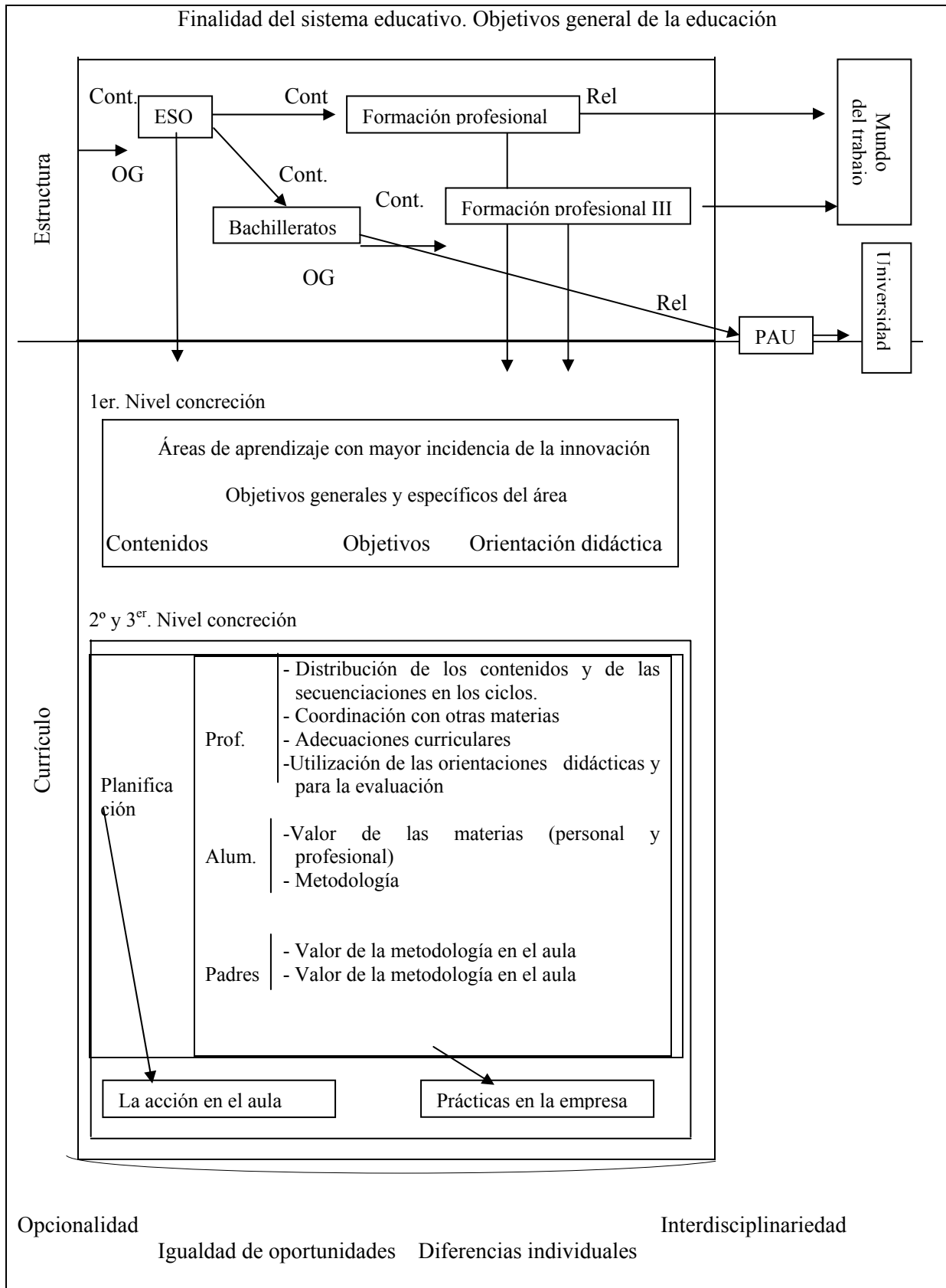


Figura 7.4. Esquema de la planificación de la evaluación del currículo.

la experiencia y de los resultados obtenidos en el momento de la evaluación. Esto implica que el grupo de trabajo del núcleo A de currículo, analiza los objetivos y la planificación inicialmente propuestos para la Reforma Educativa, los documentos escritos que hubiera, los proyectos concretos, la realidad educativa del centro y del aula así como su evolución (estrategias de aprendizaje, medios, intereses y motivación de los alumnos, comprensión, estudio, resultados, etc.). Además hace análisis comparativos entre el modelo vigente y el de cambio y busca las interrelaciones entre datos del *sector académico*, del *sector del trabajo* (proyectos curriculares de materias, desarrollo de las materias en el aula desde la perspectiva profesional, prácticas de los estudiantes en la empresa, necesidades del mundo laboral) y el *sector social* (visión del cambio desde entidades como pueden ser los ayuntamientos, desde la óptica de los padres, de corporaciones, etc.).

En el sector académico, el ámbito del currículo, se interrelaciona con todos los demás ámbitos.

Es de especial interés la proyección del modelo de evaluación EI para detectar las efectividades de la introducción del mundo laboral en la formación del alumno de los bachilleratos. La evaluación se efectúa en dos direcciones: desde el sector académico y desde el laboral. La figura 7.5 presenta el esquema de este apartado.

7.6.2. El ámbito de orientación y tutoría

Este ámbito se justifica no sólo por la gran relevancia de su función en la formación del alumnado sino también por la especial mención que hacen los planes de la Reforma Educativa.

Los contenidos se organizan en las siguientes áreas:

- Funciones y objetivos de la tutoría: estrategias y campos de intervención.
- Formación de los tutores: ámbitos de formación, intensidad, nivel de satisfacción, recursos materiales y ayuda técnica, formación futura.
- Organización de la tutoría: elección de tutores, dedicación, coordinación, programación, conexión con otros servicios.
- Conocimiento del alumno: estrategias, actividades, técnicas que se utilizan expedientes-registro.
- Contenido de la práctica tutorial: relación tutor-alumno, tutor grupo, tutor familia; opcionalidad y elección de créditos; técnicas de estudio.

La muestra de estudio escogida entre los centros experimentadores está constituida por centros públicos y privados que hayan tenido continuidad en su labor tutorial desde el inicio de la innovación, que hayan realizado la experimentación como mínimo en un ciclo completo, y que tengan conjuntamente los ciclos 14-16 y bachillerato. Debe recordarse que todos los centros que experimentan la Reforma Educativa en Cataluña no desarrollan necesariamente todos los niveles.

El equipo que realiza el estudio obtiene los datos para el análisis de los tutores, de los alumnos, de la dirección de los centros y de los coordinadores de tutores. Las técnicas utilizadas para obtenerlos son los cuestionarios (tutores, alumnos), la entrevista de grupos

(de tutores, de coordinadores con tutores y con directivos) y el análisis de contenidos de los materiales y documentos publicados por la administración y

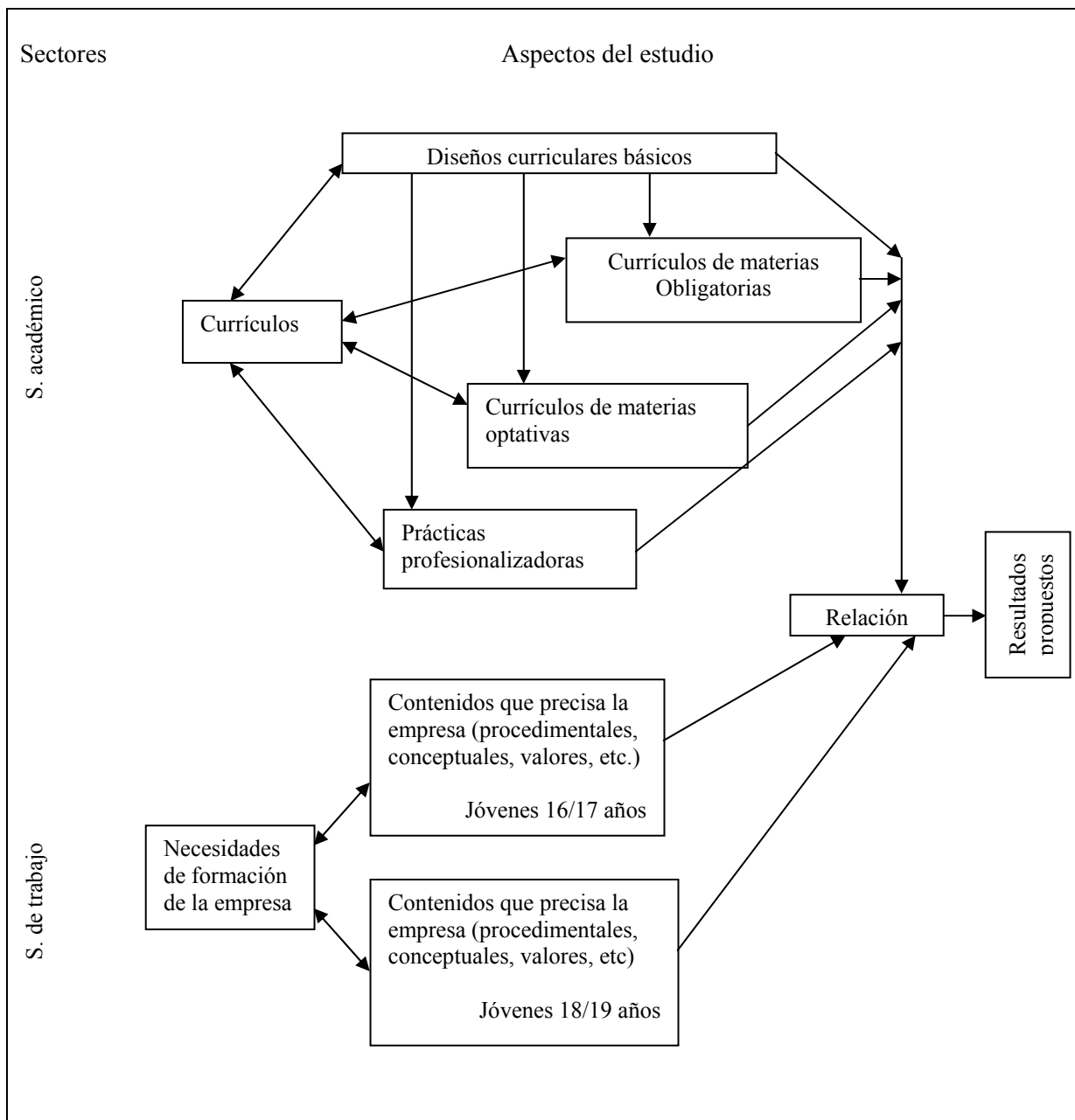


Figura 7.5. Esquema de la evaluación comparativa del contenido del currículo y de las necesidades formativas del mundo laboral

efectuados en los centros. El sector en donde se desarrolla el trabajo de evaluación es el académico y, en éste, se relaciona con el ámbito del profesorado y currículo.

7.6.3. El ámbito de la organización y gestión de los centros educativos

Los objetivos específicos de este ámbito de estudio se resumen en el conocimiento de.

- El nivel de cumplimiento, de utilidad y de suficiencia de las normas emitidas por la Administración en los centros experimentadores del cambio.
- El grado de elaboración de los planteamientos institucionales propios considerando los aspectos en que han incidido en el plan de la reforma.
- Los órganos unipersonales que hay en los centros, su proyección laboral, el grado de suficiencia, la estabilidad y su ejercicio.
- Los criterios para adscribir a los profesores por materias y por grupos de alumnos.
- Los criterios para agrupar a los alumnos y contribuir a su promoción.
- La gestión general del centro.
- La estructura *Staff* y su desarrollo en relación a la situación de la Reforma Educativa.
- El sistema relacional del centro en lo que respecta a la participación, utilidad, nivel de satisfacción, autonomía y toma de decisiones.
- El sistema de comunicación-información en relación a contenidos, procedimientos, eficacia y satisfacción.

Los contenidos a evaluar se desprenden de los objetivos mencionados y se concretan en los siguientes apartados:

- Condicionantes externos:
 - La legislación.
 - Los procedimientos complementarios para la difusión y explicación de la normativa.
- Los planteamientos institucionales
 - Planteamientos implícitos.
 - Planteamientos explícitos.
- La estructura organizativa:
 - Vertical (órganos unipersonales y colegiados).
 - Horizontal (profesores, alumnos, PAS).
 - *Staff* (base de gestión, base instructiva, base formativa, servicios).
- El sistema relacional: la participación, la comunicación, las conductas de relación son los aspectos esenciales de análisis.
- Los recursos del centro que se analizan son los materiales, personales y funcionales.

Las fuentes de información para la obtención de datos son los equipos directivos, los coordinadores de la Reforma, los profesores, los alumnos, el personal de administración y servicios, la asociación de padres del alumnado, los inspectores, y los documentos emitidos por la Administración y por los propios centros. Las técnicas utilizadas son las del

cuestionario (con preguntas abiertas y cerradas), la de entrevista (dirigidas y semicerradas), y la del análisis de contenidos para documentos.

Los estudios del grupo de trabajo se concretan en el *sector académico* y se relaciona con el ámbito del *profesorado*.

7.6.4. El ámbito del profesorado

Los autores del modelo parten de la concepción de que la formación es esencial para el desarrollo de la profesión y por consiguiente influya directamente en la calidad y en los resultados de la misma. Este punto de partida da lugar a la obtención de información sobre la formación inicial, la formación continuada o permanente recibida por el profesorado antes de intervenir en la experimentación de la Reforma Educativa o durante la misma (realizada en el propio programa de Reforma o externamente a él). También sobre la experiencia profesional. Por otra parte, en relación con los otros grupos de trabajo del sector académico, estudia la proyección del profesor en su realidad, sea en la curricular, en la de organización y gestión del centro o en la de tutoría (gruesa en la curricular, en la de organización y gestión del centro o en la tutoría (grupo de trabajo, colaboración en el centro, alumnos, aula, tutoría, etc.).

Los núcleos de contenido en el ámbito *profesorado* se concretan en:

- Historia profesional del docente: formación inicial, experiencia docente, formación permanente, acceso al centro de Reforma.
- Funcionamiento de los equipos de los centros en relación al currículo, orientación y tutoría, estructura organizativa de los propios equipos, dinámica del grupo, actividades, relaciones entre grupos de materia.
- Actuaciones de formación permanente de los docentes de centros experimentadores en los grupos de trabajo y en formación exterior a ellos.
- Organización individual del enseñante en relación al currículo y a la tutoría.
- Incidencia de los profesores en el desarrollo de los alumnos: motivación e interés, interdisciplinariedad, técnicas de estudio, procedimientos, hábitos, conceptualizaciones, etc.
- La formación del profesorado de los centros experimentadores desde la Administración: cursos, materiales, grupos de trabajo.

La figura 7.6 refleja estos aspectos así como las relaciones y dependencias existentes entre ellos.

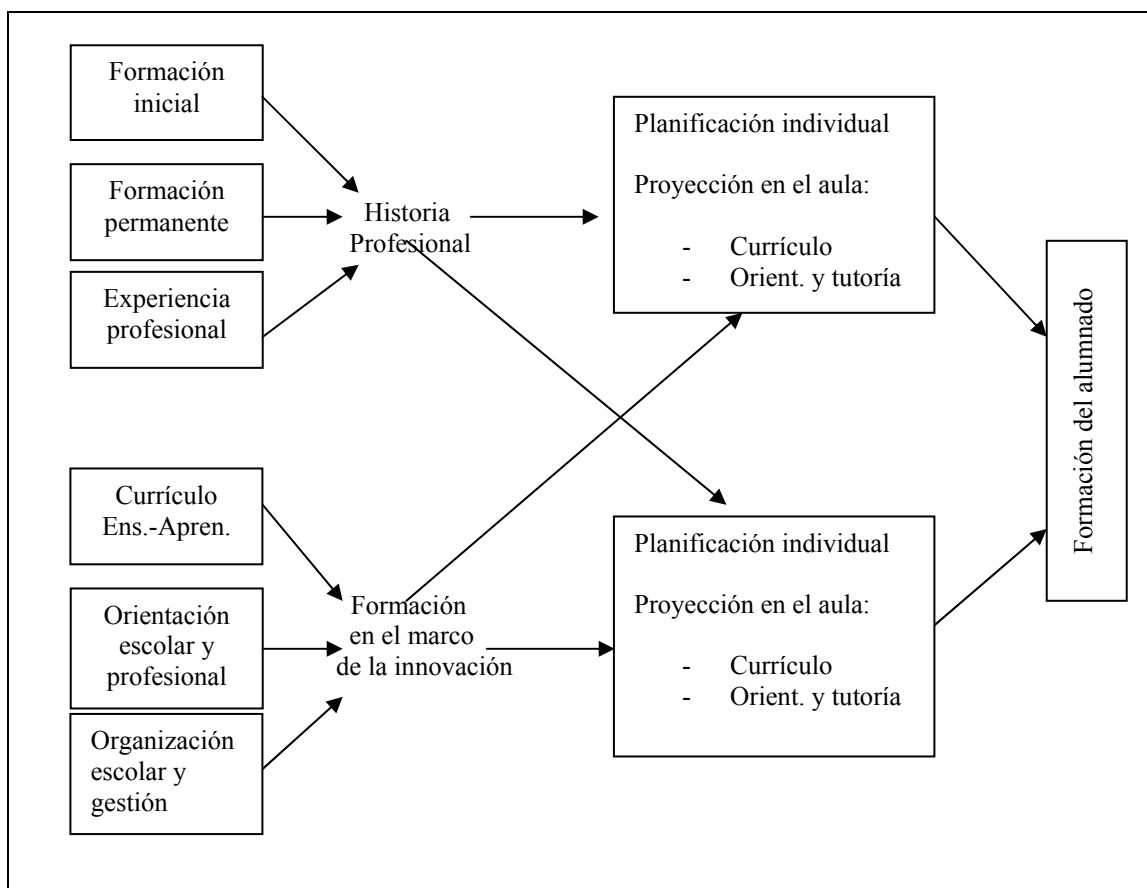


Figura 7.6. Esquema de la relación de los elementos que se consideran en la evaluación en el ámbito profesorado

Los estudios realizados en este ámbito se relacionan con los desarrollados en el *sector académico* en *currículo*, en *orientación y tutoría* y en *organización y gestión de centros*. También se obtienen datos del *sector social* (colegios profesionales, entidades públicas, padres, etc.) y del *sector profesional*.

7.6.5. el ámbito de los costos económicos

En el momento de promover una reforma, una innovación educativa, uno de los interrogantes que se impone resolver es el del cálculo de su coste sea global, sea por unidad de enseñanza, sea por alumno.

Este modelo ha enfocado su tarea considerando como base el análisis de “coste eficacia” o rendimiento efectivo de las inversiones aplicadas a la enseñanza. Eso comporta la necesidad de disponer de datos de todos aquellos aspectos específicos que de alguna manera afectan a los costos. Con ello se pretende:

- Determinar en qué grado de variables pedagógicas afectan el coste de la educación.

- Comparar la estructura pedagógica y económica de los centros con estructura homóloga y los resultados que consigan.
- Determinar los índices, las ratios y la tendencia central de cada uno de los colectivos.
- Determinar los aspectos estructurales y las variables que afectan más a los elementos óptimos y la economía de los esfuerzos pedagógicos.

Este estudio económico se efectúa con el modelo MECE (Modelo Estudio de costes de la Educación) del Ibar (1987) cuyo análisis estructural de la función educativa facilita el estudio estandarizado de las actividades académicas y su repercusión económica en los diversos centros de coste. Es un modelo de tipo administrativo aplicable a ejercicios contables cerrados que relaciona variables de tipo económico y de producción.

Los datos considerados por el grupo de trabajo son: el coste de personal (sueldo y seguros sociales), los costes de bienes y servicios, y el financiamiento (amortización e intereses de reposición) de la educación.

Los estudios muestran que en la Reforma Educativa hay diferencia entre los costos de la enseñanza pública y la privada y que los planes experimentales tienen un coste más elevado que los del plan vigente. Además el estudio presenta el costo por unidad escolar de centros en relación al número de opciones (los cálculos se realizan para centros de una a siete opciones).

El informe final presenta los resultados desde dos perspectivas: una global en relación a los objetivos generales de la evaluación y a los grandes ejes de la Reforma educativa y otra específica, pormenorizada, en la que se relacionan los aspectos diversos que afectan a normas, a orientaciones y documentos generados por la Administración y por los propios centros, a la planificación, a la realidad del centro y del aula, y a la repercusión en la formación del alumno en relación así mismo, al mundo laboral y al social.

Esquemáticamente los resultados del informe final se pueden agrupar en:

1. Repercusión de los dos aspectos básicos y de innovación que implica la Reforma Educativa: el marco curricular y la nueva estructura organizativa.
2. Los cambios que aporta la Reforma Educativa en relación a las necesidades de la sociedad y el medio, dando especial prioridad a aspectos sociales y laborales.
3. La influencia, validez y utilidad de los documentos curriculares y de trabajo desarrollados. Contenidos que se expresan, Interdisciplinariedad. Individualidad, igualdad de oportunidades, mundo del trabajo y profesionalización. Dimensión nacional y europea, conocimiento de derechos y deberes. El tercer nivel de concreción.
4. El planteamiento y proyección de la orientación educativa y profesional, así como de la acción tutorial.
5. el menor abandono escolar.
6. Los alumnos y las pruebas de acceso a la universidad.
7. La organización y la gestión de centros con la Reforma Educativa. Las normas y orientaciones de la Administración y su proyección a la realidad. Documentos elaborados por los propios centros. Introducción de cambios en los centros originados por el desarrollo de la Reforma Educativa.

8. El profesorado: formación y proyección. Formación específica en relación a la Reforma Educativa. Efectividad. Intereses. Necesidades de formación generadas por la Reforma Educativa en los diversos ámbitos.
9. El coste de la enseñanza en la Reforma Educativa: incremento y causas (mayor individualización y optatividad, necesidad de nuevos recursos y medios, mayor acercamiento al mundo laboral, etc.). Diferencias de costos entre la enseñanza pública y privada: causas.
10. La posibilidad afirmativa de la generalización de la Reforma Educativa considerando una buena planificación e información (social y a todos aquellos que deben llevarla a término) así como formación específica para los profesionales. Tener presente, además el gran esfuerzo económico que conlleva no sólo la propia Reforma Educativa sino su puesta en marcha.

Puede concluirse indicando que EI se define como un modelo abierto y de fácil utilización para distintos campos de la educación y de la formación. Además, admite ser utilizado en estudios de diferente amplitud y proyectado en su globalidad o parcialmente – un ámbito-. El modelo considera la relación innovación educativa y realidad. Su validez y fiabilidad ha sido ampliamente demostrada. Es un modelo pasivo para la evaluación de programas para la innovación y permite tomar decisiones genéricas y específicas.

ACTIVIDADES DE DESARROLLO

1. En un centro de formación continua de técnicos superiores se han introducido recursos didácticos para mejorar el rendimiento de los aprendizajes al mismo tiempo que presentar una docencia más motivadora. Los recursos específicos son: películas (presentan casos, situaciones, etc.), transparencias (para una mayor claridad en las explicaciones) y programas informáticos de simulación.
 - a) presentar más concretamente el programa de innovación, considerando algún centro que se conozca.
 - b) Indicar las funciones externas e internas de la evaluación de programa.
 - c) Explicar los elementos base del proyecto.
 - d) Realizar el proyecto de evaluación del programa para la innovación.
 - e) Determinar la organización del equipo de evaluación y el rol de cada uno de los miembros.
 - f) Explicar la realización de la evaluación e indicar quién tomaría las decisiones y por qué.
2. En un centro social de un barrio se está realizando un programa educativo para la innovación. Su finalidad es la de motivar a los jóvenes de la zona para que sepan utilizar positivamente su tiempo de ocio.
 - a) Describir más específicamente el programa para la innovación, en relación a una zona o a un centro que se conozca.
 - b) Definir el proyecto de evaluación de programa para la innovación para este caso.

- c) Hacer un análisis de lo expuesto e intentar presentar otro proyecto –desde otra perspectiva- para proceder a su evaluación.
- d) Dar una visión crítica.

8

Metaevaluación: rigor, mejora, Ética y aprendizaje

8.I. Introducción

Una buena parte de la discusión sobre los fenómenos educativos está construida con términos de una enorme complejidad semántica. La palabra educación encierra, en sí misma, tal polisemia que dentro de sus acepciones cabe todo tipo de interpretaciones y de actuaciones. Una persona pretende educar y por ello golpea al alumno o al hijo “para que aprenda. Otra persona entenderá que, para educar, no puede aconsejar por temor a la manipulación. Los dos piensan que “están educando”. Lo mismo debe decirse del concepto de evaluación. Evaluación se llama a la medición de resultados realizada mediante pruebas estandarizadas, a las auditorias, a la aplicación de pruebas objetivas, al análisis de procesos, a la investigación en la acción, etc. No digamos nada de las aplicaciones, utilizadas, usos y abusos que se llevan a cabo con la evaluación (santos Guerra, 1993). Su polisemia se multiplica hasta el infinito cuando pretendemos llegar a sus dimensiones más profundas.

Si eso sucede con los conceptos, ¿qué no pasará con los procesos que incluyen muchos de ellos? Cuando hablamos de evaluación hemos de precisar de qué evaluación se trata, a qué llamamos programa, qué es mejora de la práctica, qué entendemos por valores en la acción educativa, etc. De no hacerlo así, correremos el riesgo de estar hablando de fenómenos y procesos que nada tienen que ver con lo que otros piensan sobre ellos.

Es necesario respetar las reglas de la argumentación (Weston, 1984) para librar el discurso de trampas y sofismas. Resulta sorprender analizar los abusos lógicos que llevan a conclusiones tajantes en el ámbito de la educación y de la evaluación.

No tener en cuenta que muchos discursos, incluso lo que pretenden basarse en la investigaciones y evaluaciones más ambiciosas, están larvados por intereses y perspectivas ideológicas y políticas, es engañarse a sí mismos o engañar a otros.

No existe ninguna norma independiente. Lo primero que preguntaría al usuario un tecnólogo competente en tests de rendimiento: ¿qué porcentaje de examinados quiere que suspendan? Solo así puede proceder confiadamente el tecnólogo a construir los tests y entregar el resultado solicitado. ¿Provocó el test los errores de los interesados? [...] el test es un instrumento peligrosamente flexible cuando que queda sometido a un control político a un control político directo (MacDonald, 1977:25).

8.2. El rigor de la evaluación

Muchos discursos sobre la educación y sobre la evaluación están trucados por mor de la ambigüedad o de la peculiaridad de las acepciones de las que se parte. La secuencia “educación, evaluación, mera evaluación” va haciendo más complejo y más difícil el entendimiento y más discutible la claridad. Es interesante, al respecto, el análisis de la lógica formal de una evaluación (y de su metaevaluación) realizada por House acerca de la evaluación realizada por Glass (1972) y replicada por Scriven (1976).

El principal problema surge cuando, por creer que todos estamos hablando de lo mismo, no aclaramos lo que decimos. La delimitación conceptual es uno de los pilares sobre los que se construye el pensamiento compartido. Otro es el rigor en la argumentación, frecuentemente cargada de sofismas (Santos Guerra, 1997 a) cuando se refiere a la educación.

Como un metalenguaje, que es un lenguaje que se usa para hablar del lenguaje mismo, la metaevaluación consiste en analizar el proceso de evaluación. Digamos, sacrificando el estilo a la claridad, que metaevaluación es la evaluación de la evaluación. Aunque el prefijo “meta” puede tener connotaciones temporales (después de), no necesariamente se realiza la metaevaluación una vez finalizado el proceso de evaluación. Puede ser paralelo a la misma.

Existe una función retroalimentada de la evaluación que ha sido frecuentemente olvidada. Por eso se han repetido evaluaciones escasamente aprovechables para generar conocimiento y mejora. Por eso se han repetido errores de forma extrañamente irracional. Cuando se evalúa mucho y se mejora poco, algo está fallando en el proceso. “La función retroalimentada hace de los procesos evaluadores algo sumamente útil para conocer el valor del camino que se sigue y la calidad de la educación que se pretende” Casanova, 1992: 139).

Las ideas que se aplican a la evaluación son trasladables a la metaevaluación. No es, pues, un proceso esencialmente técnico sino que tiene naturaleza política y ética. Desde la primera y exclusiva acepción, se trataría de un proceso mecánico, aparentemente riguroso, consistente en aplicar esquemas o principios y criterios de rigor. Lo cual supondría despojar a la evaluación de sus componentes esenciales: a quien sirve, a quién trata de beneficiar.

Muchas preguntas que se formulan respecto a la evaluación han de ser planteadas en un metaanálisis coherente, concienzudo y fundamentado. Se puede aplicar un tipo de metaevaluación a una evaluación que se ha basado en criterios radicalmente diferentes. Así se verán reflejados los abusos, las contradicciones, los problemas y los desajustes.

Principalmente, una evaluación es un proceso social de aprendizaje, no sólo para las audiencias diversas del informe, sino también para los evaluadores. Por ello, metaevaluar un proceso social de aprendizaje, no puede ser menos, a su vez, que el intento de propiciar otro proceso social de aprendizaje, quizás no tan complejo como el anterior, pero no menos rico e importante. Para ello, se necesita algo más que un conjunto bien estructurado o una lista de criterios, se necesita que el metaevaluador, al menos plantee interrogaciones, sugerencias y perspectivas que alimenten la comprensión misma de los evaluadores y de la audiencia (Angulo, 1993^a: 181).

No se trata, pues, de concluir, al estilo de las investigaciones experimentales, si existe validez o no existe, sino de avivar la reflexión para generar comprensión rica y relevante sobre los programas evaluados. Esto convierte a la evaluación no tanto en una tarea de expertos cuanto en un proceso democrático de discusión.

No basta interrogarse sobre la calidad de los datos que se recogen en el informe o sobre el rigor de los métodos que se han aplicado para extraer información, sino que es preciso preguntarse por la naturaleza de la iniciativa, por la finalidad del proceso, por las reglas éticas que lo han inspirado, por la utilización que se ha hecho de la información, por el control democrático de los informes, por las cuales que la evaluación ha dejado, etc.

Quizás uno de los puntos más controvertidos y, por ello, más sujetos a discusión, e incluso a confusión, se encuentra, para cualquier perspectiva científica, en la presentación, clara y racional, de sus criterios de veracidad, es decir, de la profundidad con la que es capaz de demostrar el rigor de sus explicaciones, construcciones y conocimientos (Angulo, 1990: 95).

Plantaremos, aprovechando el concepto de contexto y algunos de los que utiliza House (1973), seis contextos básicos de exploración. El contexto permite analizar ideas en el marco concreto y temporal de su desarrollo.

Aunque todos están interconectados, se va a seguir su cronología evolutiva.

a) *Contexto de constitución*: estudia los mecanismos de generación de la evaluación, de surgimiento y negociación de todo el proceso. La historia de la génesis es muy interesante porque en ella pueden encontrarse los embriones de posteriores situaciones y actitudes. No es lo mismo una iniciativa surgida de la jerarquía e impuesta que otra que parte de los protagonistas. No es lo mismo una iniciativa que procede de los directivos de una institución que otra que parte de los integrantes de la comunidad.

En este contexto cabe incluir todo el proceso de negociación. Con quién se hace, qué contenidos aborda, qué problemática surge y cómo se trata, etc. en la negociación inicial está el mapa genético de la evaluación. Hay que diferenciar negociaciones meramente formales de otras que llegan hasta el fondo de las cuestiones. Hay que saber si se ha negociado lo insustancial dejando lo fundamental como incuestionable.

Cuestiones:

- ¿Quién y por qué ha tomado la iniciativa de hacer la evaluación?
- ¿Por qué ha surgido precisamente en ese momento y no antes o después?
- ¿qué intereses (explícitos o implícitos) determinan la respuesta en marcha de la evaluación?
- ¿Qué reacciones ha despertado el hecho de tomar la decisión?
- ¿Con quién se ha negociado la evaluación?
- ¿Qué tiempo, forma y lugar se ha destinado a la negociación de la evaluación?
- ¿Quién ha redactado el documento que servirá de vademécum durante el proceso?
- ¿Ha sido una negociación real o meramente protocolaria y formal?

- ¿Se han planteado en la negociación cuestiones fundamentales o solamente aspectos irrelevantes?
- ¿Se han efectuado cambios respecto a los planes iniciales?
- ¿Ha estado la evaluación planteada como un mecanismo de poder o de control?
- ¿Ha existido control democrático por parte de los evaluados y de los ciudadanos?
- ¿Se ha garantizado la negociación de los informes y las condiciones de su posterior publicación?
- ¿Se han explicitado las propuestas de principios éticos que van a presidir todo el proceso?
- ¿Qué reacciones ha suscitado la puesta en marcha de la evaluación?

b) Contexto de justificación: analiza los métodos elegidos para la obtención de los datos, su aplicación, la información a que han dado lugar y el tratamiento que se ha hecho de ella. Del rigor en la selección y aplicación de los métodos se deriva una buena parte del rigor de la evaluación.

Los procedimientos y estrategias utilizados para el desarrollo de la evaluación permiten conocer el “como” del proceso. La manera de proceder, la aceleración, el clima, las relaciones, las explicaciones que se ofrecen y que se piden, constituyen un elemento importante del entramado de la evaluación.

Aquí debe analizarse el tratamiento que se ha hecho de los datos, el rigor de la argumentación utilizada, la triangulación –si se ha hecho- de la información procedente de distintos sujetos y de diferentes métodos.

Cuestiones:

- ¿Qué tipo de métodos se ha utilizado para hacer la exploración?
- ¿Cómo se han realizado los registros de observación?
- ¿Se han grabado las entrevistas y se han transcrito con fidelidad?
- ¿Se ha adaptado la naturaleza y las secuencias de los métodos a las características de los interesados?
- ¿Se han utilizado métodos sensibles para captar la complejidad?
- ¿Cómo se ha aplicado cada uno de ellos?
- ¿Se ha recogido información de manera imparcial?
- ¿Se ha hecho un muestreo significativo para recoger la información de los evaluados?
- ¿Se ha diversificado la metodología para recoger una información más rica y plural?
- ¿Se ha garantizado el anonimato a los informantes?
- ¿Ha existido más de un observador para contrastar los datos recogidos?
- ¿Se han utilizado plazos temporales adecuados a la naturaleza de la exploración?
- ¿Ha existido una buena colaboración por parte de los evaluados?
- ¿Ha provocado el proceso de exploración conflictos, problemas reticencial? ¿Por qué? ¿Cómo se han afrontado?

c) Contexto de presentación: estudia el contenido y la forma de presentación de los informes, la adecuación del mismo a las diversas audiencias, la estructura lógica, la coherencia interna, la extensión, la claridad expositiva, etc.

La elaboración del informe (o de los informes) exige rigor en la argumentación que se realiza con los datos recogidos, de manera que no se produzcan silencios indebidos o conclusiones caprichosas.

La forma de presentar la información ha de tener como finalidad fundamental la comprensión de los lectores, no la epifanía de la erudición del evaluador.

Cuestiones:

- ¿Cómo se han trabajado los datos recogidos?
- ¿Tienen buena lógica los argumentos que se utilizan?
- ¿Se ha hecho triangulación de métodos, de sujetos, de tiempos?
- ¿Cómo se ha presentado la información?
- ¿Es inteligible el contenido de los informes?
- ¿Tiene el informe una buena estructura?
- ¿Tiene coherencia los datos que se presentan?
- ¿Se han hecho informes diversos para diferentes audiencias?
- ¿Han sido adecuados los plazos entre la recogida y la elaboración de informes?
- ¿El lenguaje utilizado es claro, preciso y fácilmente comprensible?
- ¿Se ha respetado el anonimato en los informes?
- ¿Ofrecen los informes un conocimiento enriquecedor de la realidad?

d) *Contexto de negociación*: el informe ha de ser entregado a sus destinatarios de forma rápida, cómoda y directa. Al decir directa se hace referencia a que tenga carácter democrático y que no esté jerarquizada. No es bueno entregar el informe sólo al director de una institución o al responsable de un programa.

Las condiciones de la negociación han de ser razonables en cuanto al tiempo y a las formas de discusión. Las actitudes del evaluador permitirán descubrir si se trata de una negociación real o de un simulacro.

Cuestiones:

- ¿Se han entregado los informes en las fechas convenidas?
- ¿Se ha concedido un tiempo suficiente para su atenta lectura?
- ¿Se ha hecho un número suficiente de copias para que dispongan de ellas los evaluados?
- ¿Se ha dedicado un tiempo suficiente a la negociación?
- ¿Ha estado presidida la negociación por un clima de respeto y de exigencia?
- ¿Se ha negociado democráticamente (o ha existido una negociación jerárquica)?
- ¿Se ha utilizado diferentes métodos de discusión (escrito/hablado, individual/colectivo, etc.)?
- ¿Se ha modificado algo del informe en aras de las aportaciones de los evaluados?
- ¿Se ha podido hablar con libertad respecto a las cuestiones problemáticas?
- Si no se ha llegado al acuerdo en alguna cuestión, ¿se ha incorporado el contenido de la discrepancia de los participantes?

- e) Contexto de valoración: analiza las motivaciones, las intenciones, las actitudes, las presiones, los cambios, las vicisitudes, etc., que han manifestado los protagonistas durante el proceso de evaluación. En el contexto de valoración se tiene en cuenta el “como” se ha realizado todo el proceso.

Al hablar de protagonistas se hace referencia fundamentalmente a tres tipos de personas: los patrocinadores de la evaluación, los evaluadores y los evaluados. La evaluación puede sufrir diversos tipos de influencias nacidas de actitudes espurias, que es necesario poner de manifiesto, ya que pueden adulterar el sentido y el significado de todo el proceso.

Cuestiones:

- ¿Se ha producido cambios respecto al plan inicial? ¿Por qué? ¿Cómo se han afrontado? ¿Quién ha tomado la decisión?
- ¿Si se ha producido algún conflicto, ¿cómo ha afrontado y solucionado?
- ¿Ha reinado un clima de confianza y de respeto durante el proceso de evaluación?
- ¿Han existido presiones por parte de algún grupo o persona que dirigir la evaluación?
- ¿Ha intentado alguien poner la evaluación al servicio de sus intereses?
- ¿Ha intentado alguien tener un acceso privilegiado a la información?
- ¿Ha querido alguien retirar la información que había ofrecido?
- ¿Se han presentado reticencias, suspicacias, suposiciones sobre la actuación del evaluador?
- Si ha habido varios evaluadores, ¿ha surgido conflictos entre sus posiciones o perspectivas?
- ¿Se ha producido algún sesgo en la utilización de la información?
- ¿Ha tenido la evaluación un ritmo adecuado en todas sus fases?

- f) Contexto de difusión: una vez negociados los informes de la evaluación ha de procederse a su difusión siguiendo unas reglas que han de estar consensuadas por los protagonistas. Es preciso tener en cuenta los derechos de los ciudadanos a conocer lo que sucede con la educación y, a la vez, el derecho de los profesionales a la privacidad en el ejercicio de su actividad.

La demora excesiva, la adulteración de los informes, la selección de alguna de sus partes, la edulcoración de los contenidos, etc., son vicios que, si existen, ha de descubrir y poner de manifiesto la metaevaluación.

Cuestiones:

- ¿Se ha hecho público el informe?
- ¿Se ha consensuado la publicación en todos los aspectos relevantes: contenido, forma, tiempo, etc.?
- ¿Han existido presiones del patrocinador o de los evaluados para eliminar alguna parte del informe?
- ¿El contenido del informe difundido ha sido claro, breve, conciso y pertinente?
- ¿Se ha recogido la opinión de los destinatarios del informe difundido?

- ¿Se ha producido algún debate de interés para la comprensión del fenómeno evaluado?
- ¿Se ha ofrecido soluciones o se han hecho aportaciones interesantes para el cambio?
- ¿Se ha aprendido algo nuevo sobre el modo de proceder en la realización de evaluaciones?

La metaevaluación es una plataforma de discusión que une a los patrocinadores, a los evaluadores, a las diversas audiencias y a la sociedad en general, que debe conocer no sólo los resultados de las evaluaciones realizadas sobre el funcionamiento de servicios públicos sino la metaevaluación que se realiza sobre ellas. La metaevaluación funciona como la garantía de la calidad de la evaluación.

La metaevaluación, como se apunta en el título, tiene que ver con diferentes cuestiones teleológicas:

- a) Con el rigor.* La finalidad esencial de la metaevaluación es el rigor de la misma. Esa cuestión tiene que ver no sólo con los métodos de exploración sino con todo el proceso que se ha desarrollado: cómo se inicia, qué negociación se realiza, qué condiciones la regulan, qué clima se genera, etc.
- b) Con la mejora.* La metaevaluación se pregunta por el resultado de la evaluación, por la capacidad de transformación de la práctica evaluada. Se pregunta por el enlace entre la evaluación y la mejora de la práctica, no sólo al final del proceso sino durante el desarrollo del mismo.
- c) Con la transferibilidad.* La metaevaluación se pregunta también por las garantías de aplicación de las conclusiones a otras experiencias y contextos.
- d) Con la ética.* La metaevaluación también ha de plantear las condiciones éticas en que se desarrolla la evaluación, la finalidad que se le atribuye, las aplicaciones que se desprenden, los posibles abusos que la pervierten, etc.
- e) Con el aprendizaje.* Hacer una metaevaluación permite aprender de los aciertos y de los errores. Reflexionar con rigor sobre la evaluación ayuda a comprender el fenómeno y permite elaborar pautas de intervención en situaciones nuevas.

La metaevaluación es una vertiente de la evaluación que cierra el proceso lógico de la reflexión sobre las prácticas profesionales. ¿Se podría hacer un nuevo bucle para evaluar la metaevaluación? Obviamente, sí. ¿Cómo saber si la metaevaluación se ha realizado con rigor? Lo mismo sucede con el meta-metalenguaje.

La metaevaluación se pregunta por el rigor. ¿Es fiable el conocimiento extraído de la evaluación o es fruto de la arbitrariedad, de los intereses y de la torpeza?

Existe un problema que consiste en trasladar los términos y conceptos de la evaluación experimental a la evaluación de carácter cualitativo. La servidumbre del paradigma etnográfico ha hecho que se haya copiado y trasladado el modelo de validez interna.

En sentido amplio –dice House (1994:231)-, entiendo por validez la cualidad de merecer reconocimiento. Ahora bien, cada modelo de evaluación hace una interpretación de lo que ha de entenderse por validez. House habla de ocho tipos de modelos de evaluación. Cuatro son calificados de “objetivistas” por su epistemología. En mayor o menor grado, todos basan su explicación de validez en métodos explícitos. *El enfoque del análisis de sistemas* afirma que es científico, que, de acuerdo con criterios explícitos, produce información fiable, datos puros y duros. *El enfoque de objetivos conductuales* afirma su validez de otra manera, basándose en la idea de la técnica. Trata de determinar la evaluación mediante la comprobación de resultados. La validez se alcanza dando cuenta de la consecución de los objetivos especificados anteriormente. No se ocupa de la naturaleza de los métodos utilizados. *El enfoque de decisión* reclama la validez en virtud de su utilidad para los responsables de las decisiones. La utilidad de los datos recogidos constituye la finalidad prioritaria, de modo que la credibilidad de la evaluación ante sus destinatarios sea alta. *La evaluación que prescinde de objetivos* del programa sostiene la validez a través de la falta de sesgos. La protección respecto a influencias inadecuadas convierte a la evaluación en un juez objetivo de las consecuencias del programa.

Las peculiaridades genéricas de estos cuatro modelos de evaluación podrían sintetizarse, según House, en las siguientes:

- a) Se derivan de la tradición teórica empírica de las ciencias sociales.
- b) Se manifiestan neutrales desde el punto de vista axiológico.
- c) Basan su pretensión de validez en métodos explícitos calificados de científicos.
- d) Se valora mucho en ellos la fiabilidad de la medida.
- e) Su fuerza está en el rigor del método prescrito, no en las interpretaciones.
- f) Se valora mucho la posibilidad de confirmar las hipótesis.

Otros cuatro enfoques son calificados por House como “subjetivistas”. *El método crítico* (similar al que ejerce la crítica de arte) consiste en la percepción holística. Establece la relación entre las partes y el conjunto de forma integrada, elaborando de este modo sus juicios. La clave de la validez reside en su habilidad para hacerlo. *El enfoque de revisión profesional* se funda en la experiencia colectiva, tradicional en la profesión. El juicio profesional, basado en las normas profesionales, hace que la evaluación merezca la atención de los destinatarios. Existen reglas, conocimientos y paradigmas, interioridades mediante las cuales puede juzgarse el ejercicio profesional. Por tanto, la validez depende del acatamiento de tales reglas de competencia aceptadas de forma intersubjetiva, aunque sea tácita. Sólo el profesional que hay participado de la cultura profesional puede haber adquirido este conocimiento y ser un juez competente. *En enfoque cuasijudicial* se fundamenta en la experiencia de la profesión jurídica, adoptando procedimientos “procesales” desarrollados en el ámbito jurídico. El grado de adaptación de la evaluación a los procedimientos adecuados constituye una medida de su validez. La clave consiste en la equidad del procedimiento. El fundamento de la validez es la experiencia judicial, tal como se estructura en la tradición jurídica. Si los procedimientos son equitativos, con independencia de los resultados, la evaluación resulta válida. *El enfoque del estudio de casos* se basa en la experiencia de los participantes y de los destinatarios. Un aspecto fundamental que se debe tener en cuenta es la comprensión por parte de los receptores. El evaluador consigna la experiencia de los participantes de manera que los destinatarios

puedan entenderla. La validez depende de la correspondencia entre la evaluación u las experiencias de los participantes y de los destinatarios.

Estas son algunas de las características fundamentales de los métodos subjetivistas en lo que se refiere a su validez:

- a) Fundan su presunta validez en la apelación a la experiencia y no al método científico.
- b) Suponen que el saber es en gran medida tácito más que explícito.
- c) Es muy importante la significación y la interpretación
- d) Se asume que cada evaluador y cada lector son capaces de establecer conclusiones de forma individual.
- e) Lo que es válido para una persona puede no serlo para otra.
- f) La evaluación ha de ser comprensible para los destinatarios.
- g) Están internamente ligados con la visión fenomenológica de la persona
- h) La exactitud de la descripción y la utilidad para los destinatarios son conceptos clave

Cada uno de los modelos tiene, como puede verse en la excelente obra de House, unas exigencias que garanticen la validez. Con sus propias palabras:

La validez de una evaluación depende de que sea veraz, creíble y correcta, desde el punto de vista normativo. Todos los enfoques de la evaluación aseguran su validez a su modo respectivo, pero la utilización de un enfoque concreto no la garantiza. Cuando más se corresponda la situación de evaluación con las premisas de un enfoque determinado, más probable será su validez (House, 1994: 236).

La tarea del metaevaluador consiste en analizar todos estos presupuestos, es decir, que tendrá que tener en cuenta el modelo de evaluación utilizado y, además deberá ir más allá para poner sobre el tapete sus limitaciones.

En cualquier caso, la evaluación debe suscitar la confianza de los interesados. Como dice House (1994:113) el evaluador, por encima de todo:

[...] debe dar pruebas de que es digno de confianza. Si el público comprende y cree las intenciones del evaluador, se reforzará la confianza en éste. La manifestación de las intenciones sirve como signo de sinceridad. Con frecuencia ésta se juzga por la coherencia de las acciones del evaluador.

La forma de pensar y de actuar del evaluador tienen que ir acompañadas de rigor en la forma de recoger la información y de consistencia en la manera de analizar los datos. No basta una intención noble y sincera. No basta un compromiso de los valores de la sociedad. Es necesario el rigor en la forma de acercarse a la realidad, de escuchar lo que “dice” y de interpretarlo de manera rigurosa y creíble.

La validez hay que encontrarla en la relación entre la interpretación y la certeza, y el proceso de justificación se caracteriza por apelar a la razonabilidad de la interpretación basada en formas de evidencia que sean públicamente accesibles (Kushner y Norris, 1990: 116).

Existen cinco modos de responder a esta cuestión, aunque no todos son de la misma naturaleza ni tienen la misma importancia para solucionar el problema epistemológico y metodológico.

8.2. I Criterios de rigor

Es necesario garantizar el rigor de la evaluación, tanto en la evaluación cuantitativa como en la cualitativa, aunque el concepto de validez tenga una diferente aceptación en un caso y otro.

Los conceptos de validez y confiabilidad son de gran importancia para los evaluadores cuantitativos que han elaborado técnicas estadísticas para asegurar la confiabilidad de los datos. Sin embargo, la validez es también un concepto importante para el evaluador cualitativo. Las técnicas para asegurarla son muy diferentes, no bien comprendidas y poco desarrolladas (Hopkins, 1989: 78).

El concepto de validez, según Hopkins (1989), consiste en que los evaluadores demuestren que las observaciones que registran y analizan corresponden a lo que pretenden registrar y analizar. La falta de validez se debe a errores sistemáticos. La fiabilidad consiste en evitar las fluctuaciones aleatorias en las observaciones que registran. La carencia de fiabilidad se produce a causa de errores aleatorios. Esencialmente la fiabilidad está relacionada con la consistencia de los resultados y se refiere al requisito de que cualquier otro evaluador, o el mismo en otra ocasión, podrá replicar la evaluación original consiguiendo datos comparables.

McCormick y James (1996) hablan de dos tipos diferentes de fiabilidad:

- a) Fiabilidad interjueces alude al grado de acuerdo entre dos evaluadores que registran sus observaciones del mismo fenómeno o de fenómenos iguales o similares.
- b) Fiabilidad intrajuez: se refiere a la consistencia de las observaciones del mismo evaluador en distintas ocasiones.

Los mismos autores hacen referencia a seis tipos diferentes de validez:

1. Validez aparente: se trata del nivel más sencillo (y débil) de validez, pues sólo requiere que una medida parezca que mide lo que se pretende medir.
2. Validez de contenido: se refiere al requisito de que los datos producidos abarquen todos los aspectos pertinentes de la cuestión de que se trate.
3. Validez referida a criterios: alude a la concordancia entre las puntuaciones de una prueba y algún otro criterio, como las estimaciones de capacidad de los profesores o los resultados de un examen.
4. Validez de constructo: la medida debe reflejar el constructo, siendo éste el conjunto de predicciones teóricas sobre una variable.
5. Validez interna: es la aceptabilidad de una explicación y de la pertinencia del diseño de investigación y los instrumentos de medida para producir la explicación.
6. Validez externa: capacidad para generalizar los hallazgos a otras situaciones.

Es evidente que cada una de estas definiciones se acomoda de forma distinta a la evaluación cuantitativa y a la cualitativa. Cuando hablamos de replicar una evaluación, está claro que nunca habrá dos situaciones idénticas si nos referimos a una exploración etnográfica. Cada experiencia es única, irrepetible, situada en un contexto peculiar, sometida a expectativas e influencias diversas, etc. Esto no quiere decir que la evaluación cualitativa tenga menos exigencias o que sea menos rigurosa que la evaluación cuantitativa. Cada una tiene su peculiaridad y su problemática.

Así como las pruebas de la validez parecen más claras y definitivas en la evaluación cuantitativa, no es menos cierto que hay que poner en cuestión el rigor de la medición en fenómenos difícilmente cuantificables y aislables.

La conducta humana nunca es estática y, como el control de ciertas variables no sería ético ni natural, ningún estudio puede replicarse exactamente, con independencia de los métodos o diseños de investigación que se utilicen. Sin embargo sigue en pie la necesidad de que los descubrimientos sean creíbles, lo que ha estimulado a los investigadores a buscar otros medios de conseguir que su trabajo sea válido y fiable (McCormick y James, 1996: 183).

La aparentemente más rigurosa forma de comprobar la validez en las evaluaciones cuantitativas no es en ningún modo absoluta e indiscutible. Es exclusivamente estocástica. Se basa en la probabilidad.

Hemos de remitirnos a Guba (1983) para plantear esta cuestión, ya que su artículo, alejado ya en el tiempo, sigue siendo un punto de referencia. Dice Guba que no existen criterios absolutos que garanticen el rigor, pero que podemos tener indicadores de su presencia.

Una de las preocupaciones más importantes respecto a cualquier evaluación son los criterios de credibilidad. ¿Tienen calidad los datos, gozan de rigor las conclusiones? Sin estas preguntas, todo el edificio se viene abajo.

El estudio de los criterios de credibilidad de la investigación cualitativa tiene que remitirse forzosamente, desde hace más de quince años, al trabajo de Guba (1983).

El contagio de la más arraigada y difundida investigación experimental ha pesado, y sigue haciéndolo, sobre las concepciones del diseño, del desarrollo y de los criterios de validez interna. Por eso se buscan criterios de parecida naturaleza (y a veces con la misma terminología) a los de la investigación cuantitativa. Es un error, ya que la tipología de los procesos es diferente.

No hay un criterio absoluto de credibilidad. Ninguno de ellos constituye una segura e inequívoca prueba de rigor. Lo cual no quiere decir que no sea necesario preguntarse por los indicadores que dan fiabilidad a la indagación. “Es dudoso que alguna vez surjan los ‘criterios perfectos’, hasta entonces, será prudente afirmar con humildad que se ha encontrado y una nueva y más verdadera senda hacia el conocimiento” (Guba, 1983: 164).

El conjunto de los métodos, ninguno de ellos absoluto e incontestable, tiene unas existencias que podríamos denominar de segundo orden, ya que se refieren a las que se deben producir entre los indicadores de credibilidad (de primer orden).

- *Contraposición*: algunos criterios tienen carácter complementario. Por ejemplo, si se toma como indicador de credibilidad la presencia prolongada del

investigador en el campo se corre el riesgo de no tener en cuenta el efecto de inmersión que sufre el investigador en el contexto. Al sentirse como nativo pierde independencia de criterio. Por eso se toma como un criterio complementario la separación periódica del campo.

- *Complementariedad*: otros criterios necesitan un nuevo indicador de credibilidad. El mismo que he citado de permanencia exige que la observación se realice durante ese tiempo esté exenta de vicios y de parcialidad. De lo contrario, mientras más tiempo se haya estado en el campo, más riesgo habrá de falta de rigor. Si se exige que se haga triangulación de la información procedente de distintos métodos y la forma de aplicar esos métodos ha sido poco rigurosa, de nada servirá que se haga un buen contraste de los datos.
- *Concatenación*: uno de los problemas que ofrecen los criterios de credibilidad es la concatenación de las exigencias. Cuando los datos iniciales están viciados, todo lo que se haga después con ellos resulta intrascendente. Es más, la apariencia de rigor de todo el proceso y de la elaboración e incluso del informe no será más que un proceso de camuflaje de la adulteración. Habrá sido igual inventarse todos los datos y luego operar con ellos de forma rigurosa.
- *Externalidad*: de la misma manera que se requieren observadores externos para que el análisis de la acción no esté condicionado por los presupuestos y los intereses de los protagonistas de la actividad, también es necesario que haya agentes externos a la investigación que acrediten (sin el sesgo que constituye el haberla diseñado y llevado a cabo) el rigor de los procesos y de los análisis.

Siguiendo los planteamientos de Guba, se revisarán a continuación los *criterios de credibilidad* que deben estar presentes en el desarrollo de la evaluación.

Métodos que se pueden utilizar durante el estudio:

1. Trabajo prolongado en el mismo lugar: la presencia prolongada no sólo pretende facilitar la obtención de datos abundantes, sino que permite “evitar las distorsiones producidas por la presencia de investigadores y proporcionar a los investigadores la oportunidad de comprobar sus propios prejuicios y percepciones” (Guba, 1983: 158).
La expresión tiempo prolongado es ambigua y aunque no ofrece ninguna referencia temporal. ¿Prolongado respecto a qué? Los externos del arco temporal no ofrecen duda, pero existe una franja intermedia de difícil valoración.
2. Observación persistente: permite “identificar cualidades perseverantes así como características atípicas” (Guba, 1983: 158). La presencia de observadores distorsiona la espontaneidad si se trata de una presencia excepcional. Cuando se convierte en asidua la presión disminuye. También se aminora la presión cuando el foco de la observación es amplio. La tercera forma de aminorar la presión que ejercen los observadores es el control democrático de la evaluación. Los protagonistas tienen constancia a través de la negociación inicial de que van a poder intervenir en la discusión del informe final.
3. Juicio crítico de compañeros: de la misma forma que los agentes externos aportan credibilidad a la acción en contextos educativos, tan influidos por las

ideas previas de los protagonistas, también las evaluaciones se benefician metodológicamente si cuentan con el juicio de expertos que no han participado en el diseño y realización del trabajo, ya que no están condicionados por la defensa de sus modos de proceder. Se podrá pensar que, si son compañeros y partícipes de la misma forma de pensar y de los mismos enfoques, será difícil que se elabore un juicio riguroso e independiente. No tiene por qué ser así. La proximidad al paradigma de evaluación hace posibles el juicio equilibrado, ya que se conocen sus presupuestos y sus exigencias. Desde la discrepancia (o la enemistad) es más difícil mantener el rigor.

Este criterio de credibilidad adquiere mayor relevancia cuando puede utilizarse durante el desarrollo de la evaluación ya que permite reorientar el trabajo. Cuando sólo se emplea una vez finalizada la investigación aporta solamente la dimensión crítica especulativa.

4. Triangulación de las informaciones: es preciso realizar la contrastación de los datos procedentes de los diferentes métodos que se han utilizado para la recogida de información. *De métodos*: las entrevistas aportan unas interpretaciones que no coinciden siempre con lo que observamos. El contenido de los documentos tiene, a veces, poco que ver con las opiniones y con la observación realizada. *De sujeto*: son numerosas las ocasiones en que las discrepancias de valoración entre las opiniones de diversas personas son sustanciales y persistentes. *De tiempo*: las divergencias de opinión son claras según el momento en que se efectúan las entrevistas. No es igual la opinión reflejada antes de celebrarse una sesión que después de finalizada la misma. *De evaluadores*: los miembros del grupo de evaluación mantienen discrepancias de interpretación en repetidas ocasiones. *De teorías*: algunos datos han de ser sometidos a la reflexión desde perspectivas teóricas diferentes, hecho que enriquece el análisis.

El tratamiento de las discrepancias y de las coincidencias ha de ser diferente según el sentido de las mismas y la importancia que tengan para la comprensión de la realidad. Algunas veces, resulta imposible resolver la discrepancia y es necesario mantener y explicar los puntos de vista divergentes. En otras ocasiones, se comprueba que algún punto de vista está movido por intereses que constituye una alteración de la realidad. En otras, se llega a un acuerdo aproximando de las versiones en una opinión integradora.

5. Recogida de material de adecuación gerencial: durante todo el proceso se ha de ir recopilando el material que acredita la recogida de la información. Las grabaciones de vídeo de las sesiones, las grabaciones de las entrevistas, las notas de campo, los protocolos de los cuestionarios, los informes de los protagonistas, los documentos utilizados, etc. todos estos materiales aportan credibilidad: en primer lugar porque han sido utilizados por los evaluadores; en segundo lugar, porque han sido objeto de análisis para el evaluador externo; en tercer lugar, porque pueden ser consultados por cualquier persona interesada en la evaluación.
6. Comprobaciones con los participantes: durante el proceso de exploración y, sobre todo, en la negociación del informe, los participantes pueden comprobar si la información que se ha obtenido es fidedigna. Por eso nos parece imprescindible favorecer la negociación de los informes de la evaluación. Es un

momento excelente de comprensión pero, también, una buena forma de interrogarse por la credibilidad de los datos. La negociación del informe es una buena forma de comprobar la credibilidad porque los interesados aceptan, rechazan o matizan el contenido de los mismos.

Métodos que se pueden usar después de concluido el estudio:

1. Establecer la corroboración o coherencia estructural: el informe ofrece una visión cohesionada y creíble de la realidad descrita e interpretada. La coherencia estructural viene a decir que las piezas del puzzle casan, que se ajustan unas a otras y que ofrecen un resultado que tiene coherencia y sentido. Lo cual no quiere decir que todas las piezas sean iguales, ni que estén colocadas de la misma forma. Las discrepancias y las contradicciones descubiertas tienen su explicación. Lo que Patton llama casos negativos o atípicos (Patton, 1984) no deben ser eliminados sino explicados. El informe tiene sentido y ofrece una visión coherente de la realidad analizada.
2. Establecer adecuación referencial: los materiales que se han recogido durante la investigación se utilizan al final para conseguir informes fundamentados, para garantizar la fidelidad de las conclusiones. Cuando se dispone de ese material, personas ajenas a la evaluación pueden comprobar que los datos y la reflexión sobre los mismos no son invención del evaluador.
3. comprobaciones con los participantes: una vez finalizada la exploración, los participantes (los grupos fuente de que habla Guba) pueden leer y analizar los datos, describir sobre ellos y discutirlos con los investigadores.

Ninguno de los criterios, considerado aisladamente, es suficiente para garantizar la credibilidad. Sin uno de ellos, por el contrario, se hace dudosa la existencia de rigor.

8.2.2. Juicio crítico de expertos

El metaevaluador se dirige a los evaluadores, pero también al patrocinador y a las diversas audiencias de la evaluación. Es importante elegir cuidadosamente al metaevaluador, ya que su independencia y las garantías de independencia que ofrece son el fundamento de un trabajo riguroso.

- a) Puede elegirse a una persona afín al paradigma de la evaluación realizada o bien a alguien que se halle situado en posiciones diferentes. Tendría dudosa eficacia elegir a alguien poco sensible al modelo de evaluación con el que se ha trabajado.
- b) La elección puede realizarla el patrocinador en convivencia con los evaluadores, o bien alguna de estas partes de forma independiente.

Es obvio que el metaevaluador realizará el trabajo desde la óptica de su peculiar visión de lo que ha de ser un proceso de evaluación.

El metaevaluador ha de disponer de toda la documentación escrita de la evaluación: documentos de negociación, informes, cuadernos de campo, etc. Y puede entrevistar a las personas que hayan participado en la evaluación.

Conviene tener en cuenta que la evaluación es todo el proceso, no sólo el informe final. Por eso es necesario seguir las pautas de revisión disponibles. O, mejor aún, comenzar el proceso de metaevaluación de forma paralela a la evaluación misma.

El metaevaluador hace juicios, pero también plantea sugerencias y preguntas (Angulo, 1993^a). Ese tipo de enfoque ayuda a los evaluadores a describir las limitaciones y a mejorar su trabajo.

El metaevaluador pone un espejo (Santos Guerra, 1995) para que los evaluadores puedan ver con más precisión cuál ha sido su manera de proceder. Mediante ese espejo, que el metaevaluador (dada su condición de independiente y dado el tiempo de que dispone para hacer el trabajo) coloca de forma precisa, el evaluador puede verse de manera más rigurosa.

8.2.3. Opinión de los protagonistas

Los protagonistas de la evaluación tienen las claves del proceso. Saben si se han respetado las reglas, conocen cómo se han aplicado los métodos, han sido los agentes de la negociación de los informes, etc.

Recabar por escrito o verbalmente la opinión de los protagonistas sobre el proceso de evaluación es un modo excelente de comprobar el rigor.

Estas opiniones sobre el proceso pueden realizarse no sólo los evaluados sino también los evaluadores. Unos y otros pueden contrastar esas opiniones que, en algún caso, se publican como complemento del informe (Soriguer, 1993; Esteva, 1993; García Arnés, 1993; Rivas, 1993; Fernández Sierra y Santos Guerra, 1993).

Las opiniones y las actitudes de los protagonistas pueden modificarse a lo largo del proceso de evaluación. Es interesante saber cómo y por qué se ha producido ese cambio.

La presencia de evaluadores externos en el programa, ha constituido para mí una invitación a reflexionar de una forma rigurosa sobre lo que estaba haciendo. En un primer momento esperaba otro tipo de colaboración. Entendía que la presencia de pedagogos consistiría en la oportunidad de desarrollar el aprendizaje de técnicas sobre la tarea que realizaba con los pacientes. Esperaba que alguien especializado en didáctica me explicase cómo podía mejorar las cosas. Me encontré, por contra, con observadores que contemplaban atentos mi trabajo, con entrevistadores que hacían preguntas (a los especialistas, a los pacientes, a los familiares) y con analistas que estudiaban los materiales que entregábamos. Después de un largo tiempo elaboraban un informe que explicaba lo que habían evaluado. Luego vi que ese era un camino novedoso, pero que conducía lejos (Rivas, 1993: 193).

Para que las opiniones sean sinceras es preciso crear un clima de libertad en el que se pueda expresar sin ambages lo que se piensa. De lo contrario, esas opiniones no tendrán ningún valor o, lo que es peor, ofrecerán pistas falsas de análisis.

Los protagonistas pueden plantear cuestiones relativas al proceso, al rigor de la aplicación de los métodos y de la utilización de los datos y, también, sobre los aprendizajes realizados:

Una de las cuestiones que más me ha llamado la atención (y también preocupado) es el importante papel que el evaluador tiene a lo largo de todo el proceso. De alguna manera el éxito del proceso evaluador va a depender no sólo de las garantías que el diseño ofrezca para garantizar, al menos, la validez interna del proyecto, sino de la personalidad del evaluador (Soriguer, 1993: 203).

No conviene distanciar mucho esta exploración a los protagonistas de la realización de la evaluación, si es que se efectúa una vez terminada, porque la separación en el tiempo hace olvidar selectivamente partes de lo que ha sucedido.

Pueden recabar esta opinión no sólo los evaluadores sino otras personas que tengan independencia respecto al proceso realizado. Ese hecho ofrecerá mayores garantías de sinceridad y, por consiguiente, de rigor.

8.2.4 Indicadores conformados

Aunque cualquier experiencia, programa, institución o actividad de carácter educativo es diferente a otra, existen unos hallazgos de la investigación que pueden servir de referentes para el contraste. Lo cual no significa que si los resultados de una evaluación difieren de lo previamente descubierto haya de considerarse ésta poco rigurosa. Sí puede servir para avivar la reflexión y revisar de forma concienzuda todo el proceso.

Respecto al soporte sobre el cual establecer los juicios de valor, algunos autores recomiendan la utilización de indicadores de rendimiento. Entienden por tales “los datos empíricos –ya sean cuantitativos o cualitativos- recogidos de forma sistemática en relación con unas metas o procesos que nos permiten estimar la productividad y/o funcionalidad de un sistema”

La formulación de los indicadores encierra una gran dificultad pero, sobre todo, el problema reside en el establecimiento de una doble correspondencia: primera, la que se pretende conseguir entre los indicadores y los programas, segunda, la que se atribuye al programa concreto con los indicadores formulados.

Se exigen unos requisitos para la aplicación de este criterio de credibilidad, a saber:

- a) Una información empírica, que aporta datos que pueden ser tanto cuantitativos como cualitativos.
- b) Recogida de forma sistemática.
- c) En relación a unos objetivos.
- d) Que permite establecer comparaciones y juicios de valor.

Si se tiene en cuenta que cada programa es único e irrepetible, será difícil establecer unos indicadores genéricos. La evaluación de los programas a través de la comprobación de los objetivos o la consecución de indicadores deja al margen preguntas de gran calado: ¿y si los objetivos estaban mal planteados?, ¿si se consiguen de forma inmoral?, ¿si se logran de manera excesivamente costosa?, ¿si se impide que los protagonistas propongan sus propios objetivos?, ¿por qué no se han conseguido los objetivos?, ¿podrían haberse planteado otros mejores (más razonables, más justos, etc.)?, ¿si se han producido efectos secundarios negativos?, ¿para qué sirven esos objetivos?, ¿han existido medios suficientes para alcanzarlos?

Otros problemas intrínsecos a la utilización de indicadores radican en los criterios de selección de los mismos, en su aplicación y en la interpretación de la información que aportan (Bottani, 1990; Selden, 1990; Sizer, 1992; Darling-Hammond, 1991; Blank, 1993).

8.2.5. El control democrático de las evaluaciones

Cuando la evaluación se convierte en un juego que comparten sigilosamente los patrocinadores y los evaluadores, existen pocas garantías de credibilidad. Por eso resulta imprescindible que los informes se hagan públicos. Son los ciudadanos los que, en definitiva, tienen el control del proceso. Conocen no sólo los contenidos de los informes sino las características que ha tenido todo el proceso.

En una democracia, los ciudadanos, los padres y los responsables públicos tienen el derecho de conocer en qué medida las escuelas públicas están cumpliendo sus responsabilidades hacia los alumnos y la comunidad social. Se necesita información increíble para que, por parte de los profesores, administración y ciudadanos, se adopten decisiones informadas (The Public Education Information Network, 1985, citado por Pérez Gómez y Gimeno, 1994).

La devolución de informes a los protagonistas, la negociación de los mismos y la difusión del informe final a los ciudadanos es una parte fundamental del control democrático de la evaluación. Esta seguridad salvaguarda a priori de muchos abusos y, a posterior, es un excelente medio para conocer y mejorar la práctica evaluadora.

Barry MacDonald habla de la evaluación democrática, lo cual supone que el control está no en manos de los patrocinadores sino en la sociedad.

La evaluación democrática constituye un servicio de información a la comunidad sobre las características de un programa educativo. Reconoce el pluralismo de valores y trata de representar un conjunto de intereses al formular los problemas. El valor básico consiste en una ciudadanía informada y su evaluador actúa como intermediario en los intercambios de información entre distintos grupos. Sus técnicas de recogida y presentación de datos deben estar al alcance de personas no especializadas [...] (MacDonald, 1974, citado por Gimeno y Pérez Gómez, 1983: 475).

Cuando los evaluadores utilizan un lenguaje indescifrable para los evaluados y para las diversas audiencias interesadas por la educación, están robando el conocimiento a los ciudadanos. La evaluación se convierte así en un mecanismo de poder que utilizan a su antojo los que pueden y los que saben.

El control democrático de la evaluación no sólo se centra en el análisis de una experiencia aislada, sino en el nivel macro de la cuestión. Los aspectos centrales sobre los que plantea el control podrían sintetizarse en los siguientes: ¿qué evaluaciones se hacen?, ¿quién decide que se realicen?, ¿por qué en ese momento preciso?, ¿qué finalidades las inspiran?, ¿qué principios éticos las guían?, ¿qué rigor tienen?, ¿a quién se encargan?, ¿cómo y con quién se negocian?, ¿para qué se utilizan?, ¿quién las conoce?, ¿qué efectos producen? ¿cuánto cuestan?, ¿quién tiene el control?

Mientras más amplia sea la extensión de los ciudadanos, más democrática será la metaevaluación. MacDonald preguntaba en una sesión de trabajo celebrada en la Universidad de Málaga en 1997: “¿Quién decide, conoce, opina y utiliza la evaluación? La respuesta a esta cuestión se aproxima a las exigencias democráticas cuando es del siguiente tipo: Todos”. En la medida que algunos privilegiados (sean éstos patrocinadores, evaluadores o informantes) tengan especial acceso a los datos o capacidad para hacer con ellos lo que deseen, se tratará de una evaluación autocrática.

La participación de los ciudadanos en las cuestiones relativas a la educación (en todas sus fases: planificación, acción evaluación, etc.) es una garantía de su buen uso. Por eso resulta importante que los evaluadores se comprometan con los patrocinadores a entregar informes a los evaluados y a otras audiencias. De no hacerlo así, el evaluador devolverá los informes al patrocinador que hará con esa información lo que considere oportuno.

8.3. La mejora

¿Para qué se realiza la evaluación? La respuesta a esta pregunta ha de situarse en la producción de conocimiento relevante que permita mejorar la bondad del programa. No se evalúa por el mero hecho de evaluar. De ser así, estaríamos ante un fenómeno de inequívoca inmoralidad. ¿Por qué desperdiciar esfuerzos, tiempo y dinero en una tarea estéril?

Lo curioso es que siendo tan importante la finalidad, pocas veces se compruebe si realmente la evaluación ha facilitado y conseguido la producción de cambios importantes.

¿Quién ha de impulsar estos cambios? En primer lugar, los protagonistas de la acción que, al conocer lo que sucede, pueden efectuar cambios en lo que están haciendo. También ha de trabajar en la producción de cambios el patrocinador de la evaluación. Algunos pueden referirse directamente a su modo de actuar, otros se conseguirán con las ayudas que deben prestar a los protagonistas.

Este carácter ascendente del cambio, rompe la línea jerárquica de la evaluación, que fácilmente podría convertirse en un ejercicio de puro control. Si se promueve una evaluación y se descubre a través de ella que existen deficiencias en la forma de administrar las escuelas, no sería bueno olvidarse de esas exigencias y subrayar solamente las exigencias de los administrados.

Una evaluación de este tipo se convertiría en un instrumento de dominio, en un mecanismo de poder que, en buena lógica, rechazarían los protagonistas de la acción. La evaluación, así planteada, interesaría a quienes gobiernan, pero sería vista como una amenaza por los profesionales.

No es fácil conocer qué cambios se producen a través de la evaluación, pero es imprescindible preocuparse por una cuestión tan básica. Los cambios pueden plantearse en diversos niveles:

- a) En las prácticas profesionales: la estructura, las condiciones, los medios, las actuaciones pueden modificarse a la luz de lo que ha descubierto la evaluación.
- b) En las actitudes: el proceso de evaluación, realizado de forma respetuosa, puede poner a los protagonistas ante preguntas y cuestiones que les ayuden a cambiar.
- c) En las concepciones: la forma de entender la práctica puede ser puesta en cuestión a través del contraste de opiniones que se alcanza con la triangulación.

- d) Las situaciones y las estructuras que las albergan: no sólo han de cambiar las concepciones, actitudes y prácticas de los protagonistas (la educación es una práctica moral), deben mejorarse también las situaciones en que éstas se desarrollan y las estructuras organizativas que las albergan (la educación es también una práctica social y política). Si, por ejemplo, se evalúa una institución educativa, no sólo es importante que quienes actúan en ella mejoren sus prácticas sino que es imprescindible preguntarse qué papel desempeña esa institución en la red de servicios sociales y qué condiciones tiene en comparación con otras que atienden a otros tipos de alumnos y de alumnas.

Hay cambios que pueden exigirse de inmediato y otros que necesitan un tiempo considerable. No se producen los cambios de manera automática sino que requieren un clima y unos medios. La forma en que se negocia, se realiza y se trabajan los informes de evaluación facilitará o dificultará el cambio deseado.

El poder de la evaluación radica en último término “en la movilización de las energías individuales” (Patton, 1978). No se evalúa por el mero hecho de evaluar. La pretensión fundamental es la transformación de las prácticas. Una de las preguntas que pueden plantearse la metaevaluación se centra, precisamente, en la potencialidad que ha tenido para promover cambios.

Los cambios no dependen exclusivamente de los evaluados. Algunos están supeditados a decisiones de los políticos y de los administradores.

Los políticos a los que preocupa que las estructuras organizativas de las escuelas apoyen la autoevaluación deliberativa y la toma de decisiones a cargo de los docentes, tienen que arbitrar procedimientos mediante los que el profesorado en conjunto pueda examinar colectivamente las relaciones entre las estructuras organizativas y sus actividades como individuos. Si se oponen a apoyar este enfoque de “abajo arriba” del cambio educativo, podemos concluir que lo que en realidad pretende la política es el control técnico de “arriba abajo” (Elliott, 1982).

Sería interesante realizar este tipo de metaevaluación tiempo después de haber finalizado la evaluación. ¿Qué efectos ha producido? ¿Qué consecuencias ha tenido? ¿Qué cambios ha promovido? ¿Qué rastros ha dejado?

Una parte de estas preguntas se puede centrar en los efectos que ha tenido la evaluación para realizar nuevas evaluaciones. Algunas se realizan de tal manera que dejan en terreno completamente arrasado para que allí se vuelva a plantar una semilla semejante.

Como hemos apuntado en otro lugar (Santos Guerra, 1994) el prisma del tiempo permite contemplar la realidad desde ángulos y en condiciones especiales. Se puede ver qué ha supuesto la prueba del tiempo, en el caso de que los cambios hayan sido fugaces o superficiales, se haya adquirido una nueva experiencia, se hayan podido generar condiciones de viabilidad, etc.

8.4 La transferibilidad y la dependencia

Una de las cuestiones fundamentales que se plantea la evaluación es la generalización de los resultados. Esta cuestión encierra tres matices diferentes.

- a) La cuestión epistemológica que plantea la naturaleza científica de los hallazgos extraídos de un solo caso. La cuestión podría sintetizarse en la pregunta que sirve de título a un trabajo de Helen Simona: *¿Existe ciencia de lo particular?*
- b) La cuestión ética de hacer público el conocimiento obtenido en la investigación, sobre todo si está financiada con fondos públicos. *¿Es legítimo privatizar el conocimiento público?*
- c) La cuestión práctica sobre la transferibilidad de los resultados. *¿Es posible aplicar a otros contextos los resultados obtenidos en la evaluación?*

La evaluación naturalista no formula generalizaciones del tipo de las que surgen de los estudios racionalistas, experimentales, cuantitativos, articulados sobre procesos estadísticos que concluyen en afirmaciones o negaciones de significatividad.

Los naturalistas, evitan las generalizaciones ya que virtualmente todos los fenómenos sociales o de conducta dependen del contexto. Creen que no es posible desarrollar enunciados verdaderos que tengan aplicabilidad general. Por el contrario, uno debe conformarse con enunciados verdaderos descriptivos o interpretativos de un contexto dado, enunciados ideográficos o relevantes para ese contexto (Guba, 1983: 150).

En la evaluación cualitativa de programas no se acaban formulando conclusiones “válidas”, con un determinado nivel de probabilidad, para todos los programas similares. Ninguna de las situaciones es exactamente igual a la otra. Las experiencias sociales no son replicables.

8.4.I Transferibilidad

Cada contexto tiene su carácter concreto e irrepetible y de su estudio no se extrae la comprobación de hipótesis de carácter general. La experiencia es irrepetible en sí misma y es difícilmente aplicable en su integridad, ya que otro programa se moverá en otras coordenadas y en otro ambiente distinto. A los investigadores cualitativos les preocupa la validez, los cuantitativos hacen hincapié en la confiabilidad y la reproducibilidad de la investigación. No es posible repetir una investigación cualitativa, ya que las condiciones son cambiantes.

Esto no quiere decir que a los investigadores cualitativos no les preocupa la precisión de sus datos. Un estudio cualitativo no es un análisis impresionista, informal, basado en una mirada superficial a un escenario de personas. Es una pieza de investigación sistemática conducida con procedimientos rigurosos, aunque no necesariamente estandarizados (Taylor y Bogdam, 1986).

La investigación cualitativa trata de explicar *esta* realidad. Ahora bien, ¿tiene alguna utilidad la evaluación cualitativa de un programa, más allá de aquello en lo que afecta al programa en cuestión? En definitiva, ¿existe ciencia de lo particular? A este proceso, en el que nos preguntamos y respondemos sobre la capacidad de aplicar los resultados de una evaluación concreta a otras situaciones de una forma rigurosa y fundamentada, le llamamos de *validación externa*.

La *transferibilidad* (Walter y otros, 1980) es un proceso que nos permite aplicar los hallazgos conseguidos a otra situación diferente/similar. Cuando nos enfrentamos a otra

situación, en parte distinta (ya que no es la misma) y en parte semejante (ya que tiene elementos comunes), nos preguntamos: ¿hasta qué punto son válidos en esta nueva situación los resultados obtenidos en la primera? ¿Son aplicables a este contexto aquellas conclusiones? ¿Se encuentran aquí los mismos procesos?

Una parte considerable de la validez externa descansa en el análisis del lector, que una vez conocida la evaluación y la realidad nueva por la que se pregunta, establece las conexiones oportunas. En este mismo sentido se manifiestan Guba y Lincoln (1982b). Igualmente al respecto se manifiesta Stake (1984), dejando en manos del lector una parte del rigor de la transferibilidad.

El primer requisito para alcanzar la validez externa es que exista un suficiente nivel de validez interna. Sin ésta, no es posible hablar con un mínimo de rigor de la primera. Es decir, que si el trabajo no tiene consistencia interna, no será posible aplicar sus hallazgos a otra realidad, sea o no similar.

Además, serán necesarios otros requisitos:

1. Similitud de contextos]: en la medida en que el nuevo contexto tenga más similitud con el primero, será más fácil que los procesos y los resultados tengan alguna relación.

Cuanto más “representativa” sea el programa, mayores serán las probabilidades de validez externa de los resultados. La transferibilidad se ve fortalecida, lo mismo que se ve fortalecida la teoría, y esto se puede hacer de diversas maneras, ya sea mediante más estudios singulares de escuelas, mediante otras formas de comprobación empírica que repercuten en la teoría, o en partes de ésta, y, con la misma importancia, en la mente de los lectores, en la medida que despliegan su propio conocimiento y experiencia de tales instituciones (Woods, 1987).

De cualquier manera, sólo hablamos de semejanzas, no de isomorfismo completo. Ninguna realidad es idéntica a otra en la esfera de lo social. No se producen leyes explicativas, causalidades perfectas que expliquen la realidad. Todas las realidades son interpretables desde un *interaccionismo simbólico* que atribuye significados sociales (Blumer, 1969) a las realidades circundantes.

2. Recoger abundantes datos descriptivos: las descripciones permiten comparar un contexto con otro. Al segundo contexto se le aplicaría la transferencia (Geertz, 1973) de lo encontrado en el primero

Esa transferencia no se realiza sobre el terminal de la explicación, sobre el *si/no* final sino sobre el proceso, sobre el funcionamiento de la *caja negra*.

La transferibilidad depende, hemos dicho, de la equivalencia de características de los centros. Y también de la especificación de las descripciones, porque lo que se quiere contrastar no son los resultados terminales sino la similitud de los procesos.

3. Las afinidades legales: el programa está regido por unas invariantes legales que, si bien serán interpretadas y adaptadas a la peculiar idiosincrasia de cada institución, son “iguales” para todos.

Si la autonomía legal de los programas fuese muy grande (el diseño curricular, en su desarrollo y control) sería menor la capacidad de transferencia de los estudios cualitativos.

Todo ello permite establecer unas condiciones de transferibilidad. Nos estamos refiriendo a realidades reguladas por un cuerpo de disposiciones idéntico que, aunque deja un margen de libertad, condiciona y dirige el desarrollo de la acción.

4. Paralelismo epistemológico: si se explicitan los supuestos epistemológicos subyacentes que llevan a presentar los procesos y los descubrimientos desde una perspectiva determinada, si se analizan con criterios rigurosos los fenómenos, si se presenta fundamentadamente la dinámica metodológica, es posible abrir el camino de la credibilidad de la evaluación y de la transferibilidad de lo descubierto a otra situación similar, con tal de que no se modifiquen sustancialmente los presupuestos citados.

De poco sirve la transferencia si al contemplar la nueva realidad escolar modificamos el punto de vista y cambiamos los presupuestos de la interpretación.

El proceso de interpretación, que actúa como intermediario entre los significados (o disposiciones a obrar de cierto modo) y a la acción misma, ha de mantenerse constante.

5. Desarrollo y verificación de la teoría: la evaluación cualitativa de programas no tiene como única misión comprender y mejorar la actividad de *esos programas*. Se pretende también verificar teorías sobre formación, organización aprendizaje, comunicación, rendimiento, etcétera y genera nuevos conocimientos.
6. Multiplicación de evaluaciones: la proliferación de estudios rigurosos de carácter cualitativo aumentará la credibilidad de los estudios convergentes en los que se confirma la bondad de la explicación de los fenómenos.

Aquí nos referimos a un tipo de inducción de carácter enumerativo, diferente al que era objeto de consideración en el epígrafe anterior. La inducción, probablemente incompleta, aportará dosis más potentes de credibilidad a medida que la explicación sea intensa y el número más amplio.

7. análisis progresivo: la evaluación de un programa permite partir, en esta profundización de los conocimientos, de presupuestos ya conocidos y contrastados. No se parte de cero cada vez que se realiza la exploración de un ecosistema escolar.

La credibilidad de los resultados de la evaluación de un programa dependerá también de su conformidad con tesis fundamentadas, respecto a las cuales supongan un avance o una profundización, una explicación más clara.

Cabe cerrar este apartado sobre la metaevaluación recordando algo que Glaser y Strauss (1967) subrayan al hablar de la credibilidad de los estudios cualitativos: son los lectores los que, en última instancia, han de juzgar sobre ella. Sería paradójico afirmar que la credibilidad de los estudios cualitativos puede fundamentarse de manera cuantitativa.

8.4.2. La dependencia

La dependencia es, según Guba, la estabilidad de los datos obtenidos en la investigación. La estabilidad está amenazada por la naturaleza cambiante de los fenómenos que se estudian en el campo de la educación y de los mismos cambios que sufren los conocimientos y estados de ánimo del investigador, que es el principal instrumento de la exploración y del análisis.

La dependencia exige la “rastreadibilidad”, es decir, la posibilidad de explicar cómo se han producido los cambios, de justificar por qué se han modificado las decisiones y por qué se ha variado la utilización de los instrumentos que se había previsto o que se estaba realizando.

La dependencia tiene que ver con la credibilidad de la evaluación, aunque no se identifica con ella ya que aquélla añade el factor temporal.

Para garantizar la dependencia de los datos obtenidos en la evaluación se pueden aplicar los criterios propuestos por Guba en el citado trabajo que, aunque él califica de primitivo, sigue teniendo vigencia.

1. Utilización de métodos superpuestos: se trata de aplicar a la vez métodos diferentes. Es un tipo de triangulación que incluye el factor temporal. La debilidad de un método se compensa con otro método. La ocultación o la deformación que puede darse en una entrevista se corrige con la observación directa de las sesiones. La estabilidad de los datos se refuerza si se consiguen los mismos datos utilizando métodos distintos. Lo que Guba llama investigación de operaciones múltiples se consigue al aplicar una diversidad de métodos que, además pueden ser aplicados por diferentes personas.
2. La réplica paso a paso: lo que importa en la dependencia es garantizar la utilización correcta del método de manera que se pueda saber si han existido anomalías o cambios injustificados. Por todo ello:
 - a) Se establecen pistas de revisión, dejando constancia del trabajo realizado y de los pasos que se han seguido para llevarlo a cabo.
 - b) Se garantiza la posibilidad de que el observador externo revise la dependencia. El observador ha de tener acceso a toda la documentación.

Es importante aprovechar las evaluaciones realizadas con el fin de mejorar las experiencias. Esto exige la difusión de los informes y que en éstos se expliciten de forma clara y rigurosa cómo se ha llevado a cabo el proceso de evaluación.

La política de difusión de informes es fundamental para que puedan conocerse y utilizarse las evaluaciones realizadas. La difusión exige canales ágiles y seguros de comunicación entre profesionales.

8.5. La ética en la evaluación

La metaevaluación no es un proceso exclusivamente técnico. Como la misma evaluación, tiene una dimensión sustancialmente ética. Por eso, la metaevaluación no se centra exclusivamente en la búsqueda de criterios de carácter metodológico o algorítmico sino que se ha de centrar en su vertiente moral.

En otro lugar (Santos Guerra, 1993) se hizo referencia a los abusos de la evaluación. Lo que en ese trabajo se describen y se denuncian han sido descubiertos en la práctica de la evaluación de nuestro país. No estamos hablando, pues, de memoria. No se hace necesario desarrollar el contenido de estos enunciados que hablan por sí solos.

- Convertir la evaluación en un elogio a quien la patrocina o la realiza.
- Elegir sesgadamente para la evaluación algunas parcelas o experiencias que favorezcan una realidad o una visión sobre la misma.
- Hacer una evaluación de diferente naturaleza y rigor para realidades igualmente importantes.
- Convertir la evaluación en un instrumento de dominación, control y opresión.
- Poner la evaluación al servicio de quienes más tienen o más pueden.
- Atribuir los resultados a causas más o menos supuestas a través de procesos atributivos arbitrarios.
- Encargar la evaluación a equipos o personas sin independencia o valor para decir la verdad.
- Silenciar los resultados de la evaluación respecto a los evaluados o a otras audiencias.
- Seleccionar aquellos aspectos que permiten tomar decisiones que apoyan la iniciativa, ideas o planteamientos del poder.
- Hacer públicas sólo aquellas partes del informe que tienen un carácter halagador para el poder.
- Descalificar la evaluación achacándole falta de rigor si los resultados no interesan.
- Dar por buenos los resultados de la evaluación a pesar de su falta de rigor, cuando esto es lo que interesa.
- Utilizar los resultados para tomar decisiones clara o subrepticamente injustas.
- Aprovechar la evaluación para hacer falsas comparaciones entre lo que es realmente incomparable.
- Atribuir los malos resultados no al desarrollo del programa sino a la torpeza, la pereza o la mala preparación o voluntad de los usuarios del mismo

Es necesario mantener la vigilancia sobre esta cuestión trascendental porque no se trata de evaluar mucho, ni si quiera de evaluar de forma técnicamente perfecta, sino de saber al servicio de quién y de qué valores se pone la evaluación. Nadie se alegraría de que se abriese una fábrica de cuchillos en una localidad si supiera que su destino era la distribución de los mismos entre los ciudadanos para ejercitarse en el acuchillamiento de sus enemigos.

El quebrantamiento de la ética puede darse en el fenómeno mismo de la evaluación, en su globalidad o en algunas de sus partes específicas. Se puede poner en marcha una evaluación al servicio del poder, el dinero o del prestigio de los patrocinadores. Por ejemplo, se puede obtener información de forma fraudulenta realizando entrevistas que están falsamente planteadas.

Care (1978) plantea las exigencias de un control equitativo de evaluación. En ese proceso se ponen los cimientos del edificio de la ética. Aunque no todas las exigencias

tienen la misma importancia ni su presentación está organizada según criterios axiológicos, es de interés enumerarlas con un brevísimo comentario.

1. Ausencia de coerción: la equidad exige que los que tienen que participar en la evaluación (evaluadores, evaluados, colaboradores, jueces, etc.) no estén sujetos a presiones coercitivas. Se sabe que esas presiones pueden ser brutales o sutiles. Internamente, tampoco deben existir mecanismos de control: nadie (persona o grupo) debe ejercer presión ni control sobre otros.
2. Racionalidad: los participantes han de actuar de forma lógica y sensata. Esto quiere decir que el discurso y el proceso de la evaluación están presididos por argumentos y no por caprichos o intereses.
3. Aceptación de los términos: las partes que firman el acuerdo aceptan libremente las reglas de procedimiento que han sido formuladas de manera clara y transparente, tanto por lo que respecta a la naturaleza, como a los objetivos y a los métodos de la evaluación.
4. Acuerdo conjunto: la equidad requiere que el acuerdo se considere como algo más que la coincidencia entre opiniones individuales. La negociación, cuando es intensa y sincera, no es un proceso sencillo. Exige tiempo. Exige valentía. Y una discusión si barreras de todos los participantes.
5. Desinterés: hay que someter los intereses particulares a las preocupaciones, problemas y necesidades de la causa común. Llegar a un acuerdo no garantiza que se hayan supeditado los intereses particulares a los generales.
6. Universalidad: la normativa tiene que afectar a todos por igual, incluidos los evaluadores. Si la normativa exige que se trate de forma distinta a diferentes personas, todos pueden situarse en cualquiera de ellas.
7. Interés comunitario: quienes llegan a un acuerdo para realizar la evaluación deben tener como meta que se produzca el máximo beneficio para cada uno de los participantes.
8. Información igual y completa: nadie debe tener acceso a una información privilegiada. Una condición de la equidad consiste en que los participantes reciban información sobre los hechos pertinentes y que ésta se distribuya por igual.
9. Ausencia de riesgos: uno de los propósitos del contrato de evaluación es que no se deriven riesgos para quienes se someten a ella, o bien reducirlos al mínimo. Scriven (1976) sugiere que se presente a los evaluadores un borrador del informe antes de entregarlo al cliente.
10. Viabilidad: el contrato de evaluación ha de ser de tal naturaleza que sea posible llevarlo a cabo. Los evaluadores no deben prometer actividades y oportunidades fantásticas, por más deseables que éstas fueran. El interés por conseguir el contrato puede llevar a realizar promesas inviables. Cuando no es posible cumplir el contrato, éste no es equitativo.
11. contar con todas las opiniones: para que el contrato sea equitativo, las partes deben tener la oportunidad de hacer constar lo que crean conveniente.
12. Participación: la condición final del contrato es que todos los participantes puedan participar. Todo el mundo debe tener voz.

Strike (1997), en una obra traducida no hace mucho en España pero que data de 1990, plantea algunos principios generales sobre la ética de la evaluación educativa.

- a) *Principio de la maximización del beneficio*: sostiene que la mejor decisión es la resultante del mayor beneficio para la mayoría de las personas. De este modo, juzga la moralidad de nuestras acciones por sus consecuencias ¿Quiénes son los beneficiarios de la evaluación? ¿Salen favorecidos por ella quienes ya lo han sido por el azar, la herencia o la cultura?
- b) *El principio de igual respeto*: requiere que respetemos el igual valor de los agentes morales. Este principio implica tres ideas subsidiarias: primero, el principio de igual respeto requiere que tratemos a la gente como fines en vez de cómo medios. No debemos tratar a los demás como si fueran simples medios para conseguir nuestros objetivos. Segundo, un hecho vital con respecto a los agentes morales es que son libre y racionales. Así pues, debemos respetar sus elecciones. Las personas libres tienen derecho a tener su propio criterio respecto de lo que es bueno para ellas. Tercero, como agentes morales, las personas tienen igual valor. Que las personas sean de igual valor significa que tienen los mismos derechos básicos y que sus intereses son de igual valor.
- c) *El principio de la intimidad*: la intimidad implica el derecho a controlar la información sobre uno mismo y protege a la gente de interferencias en sus asuntos. En la evaluación requiere que los procedimientos no sean excesivamente intrusivos y que se centre en aquellos aspectos que se refieren al contenido de la exploración.

Como se dice en el informe de la Evaluación de la Reforma del Ciclo Superior de EGB en Andalucía (Pérez Gómez y Gimeno, 1994: 20), en cuya elaboración participamos con otros evaluadores:

Se concibe como prioritaria la exigencia de garantizar la confidencialidad de los datos, mediante el anonimato más escrupuloso, de modo que deben eliminarse o camuflarse cuantos datos y elementos puedan llevar a la identificación de personas o instituciones concretas. La garantía de confidencialidad no es sólo una exigencia del propio proceso de investigación, para establecer una comunicación fluida e interesadamente generosa, sino un requisito moral indiscutible de todo proceso de producción democrática de conocimiento en Ciencias Humanas (Norris, 1998; Simona, 1987) que parte del respeto escrupuloso a los derechos del hombre, entre los que aparece el derecho a la intimidad.

No siempre es posible salvaguardar el anonimato sin quebrantar el derecho a conocer la verdad. Si solamente existe una persona que desempeña el papel de dirección, será difícil escamotear la identificación del mismo aun que obviamente, se omita el nombre del mismo. Siempre es posible encubrir a los protagonistas bajo el disfraz de los pseudónimos, pero no siempre es posible renunciar a la identificación del papel institucional de modo que se mantenga una comprensión auténtica de la realidad.

- d) *Principio de transparencia*: requiere estar abierto a las personas que tienen relación con los procedimientos, los propósitos y los resultados de la

- evaluación. El principio de transparencia protege de la jerarquización, del secretismo y de la manipulación de la información.
- e) *Principio de humanidad*: exige que se muestre consideración por los sentimientos y los individuos en contextos evaluativos. Lo decían Taylor y Bogdam de forma lapidaria: en la evaluación hemos de ser honestos, no crueles.
 - f) *Principio de no discriminación*: la igualdad, en el contexto de la evaluación, puede entenderse mejor como una prohibición contra la toma de decisiones basada en datos irrelevantes. Características como la raza, la religión, el sexo o el grupo étnico al que se pertenece se considerarán irrelevantes.
 - g) *Principio del respeto a la autonomía*: los profesionales tienen derecho a una libertad razonable en el ejercicio de su trabajo. Los evaluadores no deberían limitar ni la libertad ni el juicio.

La ética no ha de presidir la forma de llevar a cabo el proceso sino que se preocupa por el respeto a los valores en el contenido de la evaluación. No sólo de los valores que existen en el programa o la institución que se evalúa, sino de los valores de una sociedad democrática. Pondremos un ejemplo: si hacemos la evaluación de un Hospital y encontramos en él magníficos planteamientos en el trato a los pacientes, extraordinaria eficacia en los diagnósticos y en las intervenciones terapéuticas, una buena estructura de espacios y una excelente coordinación y funcionamiento de los profesionales, ¿no deberíamos decir nada sobre el hecho de que solamente pueden acceder a ese hospital los ciudadanos con altísimo nivel económico mientras a pocos metros enferman y mueren las personas que no tiene medios económicos?

La autoridad legítima sobre el proceso de evaluación proviene del modelo normativo de la organización de la enseñanza en la sociedad. El modelo condiciona la institucionalización de la enseñanza y promueve (o admite) un modelo de evaluación coherente, que no la contradiga. Por eso es necesaria la reflexión ética, ya que la evaluación puede convertirse en un elemento de transformación y de mejora del modelo hegemónico. La evaluación ha de poner de manifiesto las limitaciones éticas de los modelos imperantes para hacerlo avanzar hacia otras fórmulas más positivas.

Los valores de la visión comunitaria de la democracia son que permite altos niveles de participación en la toma de decisiones de aquellos más directamente afectados por las mismas, que éstas se toman de un modo dialogante, y que se tiende a no crear estructuras jerárquicas (Strike, 1997: 498).

La educación no es un fenómeno aséptico. La forma de organizar las escuelas tiene mucho que ver con la justicia social. Tampoco lo es la evaluación de los procesos educativos, sea cualquiera el ámbito y la temporalidad de los mismos.

Nuestras formas institucionales también reflejan nuestros valores políticos. Por ello, las propuestas de cambio de las formas institucionales actuales deben ser evaluadas a la luz de estos valores. Las tensiones existentes dentro de las estructuras gubernativas de las instituciones educativas pueden explorarse muy productivamente a partir de las tensiones entre estos valores políticos. Esta exploración puede ser reveladora con respecto al modo en que está institucionalizada la evaluación. (Strike, 1997: 485).

La ética que se exige a la evaluación ha de presidir también la metaevaluación. ¿Hasta dónde ha de llegar esta exigencia? ¿Quién evalúa a los metaevaluadores? Si se establecieran comisiones de garantía de la ética, ¿quién ejercería el control democrático sobre ellas?

En su excelente y ya clásica obra *Evaluation With Validity*, traducida al castellano con el título *Evaluación, ética y poder*, House (1980) exige a la evaluación veracidad, belleza y justicia.

Según Rawls (1971) hay que entender la justicia como equidad, lo cual no significa distribuir los recursos dando a todos lo mismo. Los principios específicos de la justicia dependerán de estos dos principios fundamentales:

- Primer principio: toda persona ha de tener igual derecho a los más amplios sistemas de libertades básicas iguales, compatible con un sistema semejante de libertad para todos.
- Segundo principio: las desigualdades sociales y económicas han de organizarse de manera que beneficien a los menos agraciados, salvando el principio de justicia y estén subordinadas a servicios y empleos abiertos a todos en condiciones de equitativa igualdad de oportunidades.

No siempre se plantea la evaluación siguiendo estos criterios. Cuando se pretende medir los resultados de las experiencias y de los centros escolares a través de pruebas estandarizadas con el fin de hacer una clasificación según la posición en la escala y se conceden luego recursos siguiendo esa clasificación, se está utilizando la evaluación al servicio de la injusticia y de la desigualdad. Por muy rigurosas y precisas que sean las pruebas, no se evita el desajuste moral. Se está dando a los que más tienen (cultura, dinero, poder, etc.) todavía más a través de un instrumento que parece neutral y aséptico. Es una trampa terrible que algunos defienden de forma ingenua y muchos de manera interesada.

8.6. El aprendizaje en la evaluación

La metaevaluación “implica que comprendamos cómo aprender las personas a partir de la actividad de evaluación. Después de todo, sólo en el acto del aprendizaje se garantiza el valor de la evaluación” (McCormick y James, 1996: 309).

La evaluación permite construir un conocimiento que se cimienta en los datos recogidos durante la exploración y en trabajo que los evaluadores realizan con los mismos. Aunque la finalidad fundamental sea la mejora, lo cierto es que la evaluación aporta un caudal de nuevos conocimientos al acervo de saber sobre la educación.

Una fuente de información y de conocimiento es el proceso mismo de la evaluación. Se aprende, a través de la evaluación, cómo se debe plantear, negociar, explorar, redactar informes aplicar resultados, etc. Resulta sorprendente la facilidad con que se repiten errores o dejan de utilizarse conocimientos ya contrastados sobre cómo se ha proceder al impulsar al impulsar la política de evaluación o al desarrollar las evaluaciones concretas. Veamos quién puede realizar el aprendizaje:

- a) Los patrocinadores: la metaevaluación puede ayudar a descubrir a los patrocinadores las limitaciones y los errores del planteamiento que han

hecho. El aprendizaje puede referirse al lanzamiento de la evaluación, a las personas a quienes se ha encargado la misma, a la utilización que se ha hecho de los informes, a la difusión de los mismos, etc.

- b) Los evaluados: además de lo que pueden aprender en el proceso a través de las preguntas, de los debates con los evaluadores, de las actitudes de los colegas, pueden aprender del contenido de los informes y de la negociación de los mismos. También puede ser fuente de aprendizaje el proceso de transformación de la práctica al que ha dado lugar la evaluación.
- c) Los evaluadores: en su relación con los patrocinadores y los evaluados, con las audiencias de la evaluación, con otros evaluadores y metaevaluadores pueden aprender mucho tanto respecto al contenido como a la metodología y a la validez de las evaluaciones. La experiencia no produce automáticamente el aprendizaje. Pero sí se puede convertir la experiencia en sabiduría cuando se reflexiona rigurosamente.
- d) Los ciudadanos en general: la evaluación no es una cuestión que monopolizan los expertos. Tampoco afecta solamente a los evaluadores cuando el objeto de la evaluación es un fenómeno tan decisivo para la sociedad como la educación.
- e) Los teóricos: la realización de evaluación y su multiplicación ha de convertirse en una plataforma de discusión para los especialistas. La metaevaluación ayudará a profundizar el conocimiento sobre la validez, la ética y la utilidad de la evaluación.

¿Cómo puede realizarse este aprendizaje? En primer lugar, es necesaria una actitud de curiosidad y de incertidumbre que mantenga despierto el deseo de aprender. La certeza condena a la esterilidad. En segundo lugar es preciso que se tenga acceso a la información, no sólo a los contenidos de los informes sino a los análisis de proceso.

El capítulo XI de la obra *Evaluación del currículum en los centros escolares*, de McCormick y James (1996:308) se titula “Aprender de la evaluación”. En dicho capítulo, estos autores (o, mejor dicho, autor y autora), hacen hincapié en la importancia del aprendizaje durante el proceso de evaluación y no solamente en el momento final de la misma:

Queremos considerar cómo puede producirse el aprendizaje a través del proceso evaluador (lo que aparece en cursiva es importante porque, en la actualidad, no disponemos de pruebas concluyentes, sólo de hipótesis). Al hacerlo así, pretendemos romper la pauta línea impuesta por la palabra escrita y, a la vez, hacer hincapié en que la evaluación curricular queda mejor conceptualizada como proceso cíclico (o espiral) que como un acontecimiento dividido en fases.

La espiral de la comprensión se va construyendo a medida que avanza el proceso y no sólo se apoya en la documentación escrita. Es más, no sólo parte de los aciertos y de los descubrimientos, sino que también puede surgir de los errores y de las limitaciones.

ACTIVIDADES DE DESARROLLO

1. Después de leer detenidamente los seis contextos básicos de exploración que utiliza House (1973) y que se plantean al principio de este capítulo:
 - a) Diseñar un gráfico en el que se plasmen las principales características y las posibles interconexiones de todos ellos
 - b) Añadir tres preguntas nuevas al listado de cuestiones que se detallan al final de cada uno de ellos.
2. Elegir un proceso concreto de evaluación y leer los interrogantes que se plantean en el apartado 8.24 (“Indicadores confirmados”)
 - a) Individualmente, ir respondiendo a estas cuestiones por escrito.
 - b) A nivel de grupo, hacer una puesta en común y elegir uno de los procesos comentados para trabajarlo en profundidad.
 - c) Discutir los diferentes aspectos que vayan surgiendo del análisis grupal.
 - d) Individualmente, realizar un informe con las conclusiones.

Elaborar un cuadro comparando los dos grandes grupos de modelos o tipos de valuación (objetivistas y subjetivistas) de los que se habla en el apartado 8.2 (“El rigor de la evaluación”).

3. Partiendo del capítulo y de la bibliografía complementaria que se considera oportuna, confeccionar un mapa conceptual del término *metaevaluación*.

BIOQUE III

"PARADIGMAS Y TÉCNICAS DE EVALUACIÓN"

DÍAZ, Barriga Ángel; “Tesis para una teoría de la evaluación y sus derivaciones en la docencia” En: Didáctica y currículum. México Ed. Paidós pp 147-197.

Capítulo 5

TESIS PARA UNA TEORÍA DE LA EVALUACIÓN Y DE SUS DERIVACIONES EN LA DOCENCIA

Cuando estudiamos el problema de la evaluación escolar, en la mayoría de los casos encontramos que existe una continuidad lineal en la manera en que lo abordan los distintos autores. Si entre los libros publicados en la década pasada nos propusiéramos buscar las diferencias en el tratamiento que se da a esta problemática, lo que encontraríamos sería una gran similitud en los temas abordados y en la respuesta técnica que se les da.

Podríamos preguntarnos: ¿Cuál es la diferencia entre los libros de Gronlund, *Elaboración de tests de aprovechamiento y Medición y evaluación*, y el trabajo de

*Thorndike y Hagen, Test y técnicas de medición en psicología y educación?*¹ ¿Cuál es la diferencia entre los textos de estos autores y los publicados por especialistas mexicanos, por ejemplo, el de Fernando Carreño, *Enfoques y principios teóricos de la evaluación*, o el de Fernando García, *Paquete de autoenseñanza de evaluación del aprovechamiento escolar?*

Podemos afirmar que si bien existen algunas diferencias entre estos textos, tales diferencias no reflejan una distinta concepción de la educación, del aprendizaje, del hombre, y un ejemplo muy claro es el trabajo del Bloom y sus colaboradores, *Evaluación del aprendizaje* que, en el proceso de modernización del discurso, llega a ser la base de su sofisticación cuando desemboca en el empleo de sistemas automatizados —computadoras— para su realización.

En el fondo, el problema de la evaluación no podrá ser visto y analizado con otro enfoque mientras subsista el mismo paradigma epistemológico para su análisis. Es necesario intentar «una ruptura epistemológica» con el trabajo realizado hasta el momento en la evaluación escolar, así como no sólo revisar las premisas epistemológicas de este discurso, sino construir un planteamiento teórico-epistemológico distinto, que se fundamente en otra concepción de aprendizaje, hombre y sociedad.

Por lo mismo, presento algunas tesis² sobre la evaluación como punto de partida para replantear su problemática. Estas tesis representan como construcciones que apuntan hacia una teoría de la evaluación y que a la vez pueden considerarse como base para iniciar el análisis y la reflexión sobre este objeto de estudio.

No intento que estas tesis tengan validez empírica, puesto que es necesario romper precisamente con el paradigma que obliga a que cualquier objeto de estudio deba ser tratado únicamente mediante el método experimental. Tampoco sostengo una postura teoricista que intente expresar que solamente por la reflexión se conoce la realidad, pues acepto que hay necesidad de concebir ésta como un mundo aparente, lleno de significados que deben descodificarse para construir la esencia del fenómeno. En el proceso de conocimiento se impone el contacto con la realidad. Así, la práctica de la evaluación escolar, como práctica social, es punto de partida y de llegada en el proceso de la evaluación. Los trabajos teóricos sobre esta temática tienen por lo tanto que ser construcciones que den cuenta de este objeto de estudio y que permitan a la vez su transformación; construcciones teóricas por medio de las cuales el hombre, en su proceso de conocimiento, intente construir la esencia de un fenómeno sólo así la evaluación puede convertirse en un «concreto de pensamiento» —parafraseando a Marx—, al ser «un concreto real»³.

La segunda parte de este capítulo plantea un conjunto de derivaciones dirigidas a la actividad docente, a partir de las premisas construidas; conjunto que adquiere sentido y valor a partir de las tesis establecidas y no por sí mismo. No pretendemos que estas derivaciones constituyan exclusivamente una nueva técnica que haya que seguir, pues sólo

¹ los títulos de las versiones originales son los siguientes: N. Gronlund, *Constructing Achievement Test*; N. Gronlund, *Measurement and Evaluation in Teaching*. Hagen y Thorndike, *Measurement and Evaluation in Psychology and Education*.

² La noción de tesis conserva dos significados: (a) designa lo que en el comienzo de una discusión el interlocutor pone como asunción propia, (b) designa una proposición tomada como principio propio. Véase N, Abbagnano 1972

³ «Los concretos es concreto porque es la síntesis de múltiples determinaciones; por lo tanto unidad de lo diverso. Aparece en el pensamiento como proceso de síntesis, como resultado, no como punto de partida, y, en consecuencia, es el punto de partida también de la intuición y de la representación» C. Marx 1977, p. 58.

son un intento de replantear la orientación de los cambios en relación con la evaluación escolar. Por tanto, tampoco se presentan como un espacio acabado y cerrado en sí mismo, sino como un espacio necesario para continuar la reflexión teórica.

TESIS PARA UNA TEORÍA DE LA EVALUACIÓN

En este apartado desarrollamos dos tesis que, como postulados, no permitan hacer un conjunto de reflexiones teóricas sobre la evaluación. Destacaremos la necesidad de abordar este objeto de estudio en una dimensión que permita hacer una construcción del mismo; por ello, consideramos que el papel de la teoría es irremplazable en este proceso, y particularmente importante en relación con una teoría de la evaluación, dado que este aspecto ha sido descuidado hasta ahora.

Primera tesis: la evaluación es una actividad social

En esta proposición subyacen dos problemas fundamentales: (1) el objeto de evaluación está inserto en lo social, por lo tanto, debe ser tratado en las llamadas ciencias humanas, y (2) la evaluación es una actividad socialmente determinada.

En relación con el primer punto, la evaluación en las ciencias del hombre, si analizamos la mayoría de los textos sobre evaluación encontraremos que en ellos no se lleva a cabo una reflexión teórica sobre su objeto de estudio, el conocimiento del proceso de aprendizaje en los hombres. Prácticamente la mayoría de estas publicaciones dedica muy poco espacio a este aspecto, para entrar de inmediato al problema de los instrumentos. En este sentido, se trata de versiones tecnicistas de la evaluación. Nuestra preocupación radica en que estos materiales privilegian el problema de las técnicas aplicadas. Sería importante, sin embargo, intentar una desestructuración de la posición tecnicista de la evaluación, reflexionando sobre la manera como terminan por otorgar un lugar clave a las técnicas y sobre los puntos epistemológicos que subyacen en sus planteamientos.

Uno de los fundamentos psicológicos de la evaluación es el discurso conductista.⁴ A principios de siglo, la psicología conductista, en su intento por adquirir el estatus de ciencia, retomó la lógica científica dominante en las ciencias naturales —la física— lo que la llevó a efectuar una reducción de su objeto de estudio, el hombre, restringiéndolo a los fenómenos conductuales observables en el sujeto. Así, manteniéndose fiel a la tradición positivista que sólo reconoce como científico lo que puede ser verificable no sólo a través de la observación empírica, sino por la posibilidad de controlar la producción del

⁴ Al hablar de «científico» para estudiar desde la psicología al hombre, aun cuando las necesidades de la misma la lleven a expresar que estudia la «conducta observable». Uno de los generadores de esta concepción, Watson, sostuvo que para alcanzar la cientificidad de la psicología «era preciso transformarla en una ciencia natural [...] Así, el conductismo es una ciencia natural que se arroga todo el campo de las adaptaciones humanas. Su compañera más íntima es la fisiología.» Posteriormente Skinner reconoce un conductismo viejo y uno nuevo; sin embargo, las premisas empiristas que sostienen este discurso permanecen inalteradas. Por ello, afirma que «el conductismo metodológico podría ser concebido como una versión psicológica del positivismo lógico». En el campo del comportamiento surge la posibilidad de que las contingencias del refuerzo puedan explicar una obra de arte o la solución de un problema matemático o científico sin recurrir a una clase diferente de mente creadora, un rasgo de creatividad. «Probablemente los científicos del comportamiento estarían muy contentos con el grado de rigor que muestra la física.» Véase J. Watson 1972 y B. F. Skinner 1975.

fenómeno, utilizó como instrumento metodológico la cuantificación del mismo. De esta manera, el objeto de la psicología tradicional, la conciencia, queda desplazado y se considera metafísica. A la vez, el método de la introspección es reemplazado por el método experimental y por el paradigma empirista de las ciencias naturales. Así, la física se convierte en el modelo de hacer ciencia, que es acríticamente imitado por algunos planteamientos de las ciencias sociales y en particular por la psicología conductista. Históricamente, el modelo de ciencia que ha ganado crédito para obtener el reconocimiento de «científico» en toda investigación, es el modelo empírico-analítico que se ha desarrollado para las ciencias de la naturaleza, éste se remonta hasta el mismo Bacon y, con la división y clasificación de las ciencias positivistas que realizó Augusto Comte, se ha venido desarrollando hasta nuestros días. Habermas⁵ considera que los métodos empírico-analíticos se caracterizan «porque sólo toleran un tipo de experiencia definida por ellos mismos, la observación controlada que requiere la organización de un campo aislado de circunstancias reproducibles y la pretensión de obtener no sólo hipótesis lógicamente correctas por la vía deductiva, sino empíricamente certeras».

Por ello, Bleger expresa que las ciencias de la naturaleza sólo logran su exactitud, objetividad, verificación, etc., amputando el «contexto y el carácter de los fenómenos y de su investigación y, por lo tanto, son métodos válidos para contextos incorrectos o limitados, creados *ad hoc*»⁶. Es así como el conductismo se desarrolló bajo la premisa de atenerse exclusivamente a los hechos observables, renunciando a todo aquello que no fuera registrable o medible; por ello se puede afirmar que:

El científico restringe su conocimiento al estudio de lo que puede observar y manipular. Su procedimiento general consiste en manipular de una manera observable cualquier condición que suponga importante en su problema, y observar luego los cambios que ocurren como consecuencia de ello. Estos cambios en el fenómeno, los relaciona con su manipulación de las condiciones como interacciones.

Por lo tanto, el científico tiene la necesidad de tomar en cuenta aquellas variables, dependientes e independientes, cuyos indicadores puedan ser observables; de esta manera el método que se pretende seguir para la investigación de los fenómenos psicológicos los desliga, los desprende de la totalidad,⁸ en tanto que se dan en una estructura social y humana que no se puede disociar; por ello «el método del conductismo arrastra un serio lastre: la renuncia expresa a investigar todo lo que no es observable»⁹. Se puede decir, siguiendo a Bleger, que no es una coincidencia que los métodos empírico-analíticos dejen de lado la totalidad del fenómeno por no ser observable, ya que:

[La exactitud que] han logrado las ciencias de la naturaleza [la han obtenido] restringiendo su campo a las necesidades y exigencias de su método: el método, así, ha restringido los objetos a los que se aplica, o mejor dicho ha construido sus objetos de acuerdo con el

⁵ Habermas 1993.

⁶ J. Bleger 1977, p. 114

⁸ Utilizo la palabra totalidad para destacar la necesaria vinculación de los fenómenos a partir de la cual éstos son comprensibles; en cierto sentido, esto se opone a la fragmentación como elemento central para el estudio de un fenómeno.

⁹ J. Azcoaga 1977, p. 95.

método. No se trata entonces de que las ciencias del hombre no entren en el método científico: se trata de que este último no entra en las ciencias del hombre.¹⁰

Habría que preguntarse entonces hasta dónde es posible trabajar con otro método más adecuado a nuestro objeto de estudio, o sea, la evaluación; un método que considere al alumno como una totalidad, que tenga en cuenta el proceso de aprendizaje y que, sobre todo, parta de la aceptación y reconocimiento de la dificultad para estudiar fenómenos complejos por las sobre determinaciones¹¹ que tienen en sí mismos, como el hombre, el aprendizaje y el proceso grupal. Dicho de otra manera, destacar la dimensión social de la evaluación permite reubicar su objeto de estudio; no se trata de adecuarlo a los requerimientos del método experimental que el discurso actual de la evaluación toma de la psicología conductista, sino de trabajar con una lógica diferente, que abra otras alternativas par ala producción de conocimientos sobre la evaluación.

A este respecto, Vasconi expresa:

Los investigadores realmente originales no sólo produjeron conocimientos nuevos sino que también produjeron nuevos métodos para la obtención de conocimientos. Pero cuando estos nuevos métodos se legitiman y se convierten en el único modo de hacer algo «científico», lo que se hace es dificultar una producción de conocimientos radicalmente nuevos.¹²

En la actualidad, nuestro reto consiste en construir estos conocimientos y esta lógica, frente a lo que Varsavsky llama el «panorama desolador de las ciencias humanas, en el que predomina la imitación acrítica de los métodos de las ciencias físicas».¹³ Por lo tanto, el discurso actual de la evaluación tiene un fuerte «obstáculo epistemológico» que refleja las premisas teóricas positivistas y pragmatistas que lo apoyan. De ahí que, en su replanteamiento metodológico, exige que se deje de abordar el fenómeno de la evaluación como fenómeno natural a la manera de las ciencias naturales, y que se le reconozca su dimensión social; que construya sus planteamientos metodológicos dentro de las ciencias humanas, lo que implica a la vez la necesidad de replantear sus propias premisas teóricas, para permitir una construcción diferente del discurso evaluativo.

Si nos preguntáramos ahora a qué se debe que la evaluación se conserve como un discurso tecnicista, vinculado únicamente al conductismo, encontraríamos que la respuesta nos lleva al segundo problema que subyace en este planteamiento: la evaluación está socialmente determinada. De alguna manera se puede hablar de que la evaluación está condicionada socialmente, y a la vez, por sus resultados, condiciona a la sociedad, ya que estos reflejan las posibilidades económicas que tienen los individuos, sus certificados de estudio, problemas de calificaciones bajas, la reprobación, etc.; también se pueden explicar por factores socioeconómicos y culturales que rodean a los estudiantes, y no sólo como un problema de falta de capacidades.

¹⁰ J. Bleger 1977, p. 114

¹¹ Empleamos el término sobredeterminación, para hablar de una causalidad múltiple de los fenómenos. Esta policausalidad es muy compleja porque deriva de distintos contextos y de múltiples relaciones. Véase J. Bleger 1977, p. 162

¹² T. Vasconi 1978, p. 8

¹³ O. Varsavsky 1975, p. 32.

Varios autores reconocen ahora los efectos sociales de la evaluación y por lo tanto afirman que esta actividad no puede ser concebida como una acción neutra, técnica y aséptica,¹⁴ sino que obedece a determinaciones sociales. Por un lado, cuando los docentes realizan esta actividad, en la mayoría de los casos no son conscientes de las repercusiones que la evaluación provoca, tanto en los individuos que son objeto de ella como en la misma sociedad. Por medio de la acción evaluativa, el maestro determina quiénes han aprobado el curso y les confiere un estatus diferente, ya que por esos resultados serán considerados por sus compañeros, sus padres y los demás maestros como buenos alumnos, mientras que los demás serán considerados deficientes.

De ahí que uno de los elementos centrales de la práctica educativa consiste en fomentar el individualismo y la competencia como condición básica para poder triunfar en la escuela y en la sociedad. No podemos dejar de señalar que estas acciones que realiza la escuela, por medio de sus representantes, son determinantes en la formación de los escolares. Es aquí donde se ubica el «currículum oculto» que, precisamente por su falta de explicitación, forma parte de una serie de aprendizajes latentes, más importantes que los conocimientos que la escuela enseña y determinantes en el aprendizaje del estudiante.

Vale la pena insistir en que la evaluación contribuye de alguna manera a que el estudiante pierda o no tome conciencia de sí mismo y de su propia situación, dado que lo importante es sobresalir, ganar a los otros y obtener la máxima puntuación. De esta forma, el alumno difícilmente reflexiona sobre su aprendizaje, es decir, sobre para qué aprendió y cómo logró aprender; también crea un falso mito sobre el aprendizaje, referido básicamente a un número: la máxima calificación significa que ha aprendido, la calificación no aprobatoria quiere decir que no sabe.

Por otro lado, «la misma sociedad refuerza estas deficiencias al convertirse en consumidora de las calificaciones, las cuales dominan en el sistema educativo y son usadas por patrones y empleadores como criterios para la selección de su fuerza de trabajo».¹⁵ Los certificados que expide la escuela no dejan de tener un valor social, por el estatus que adquiere el individuo que puede demostrar cierto nivel de escolaridad frente a otros que no pueden hacerlo.

Como lo señala Seve,¹⁶ no se pueden reducir los éxitos o fracasos escolares de los alumnos a un problema de «aptitudes», ya que a veces se formulan explicaciones biológicas de este problema, cuando el contexto social es determinante en el desarrollo de las mismas aptitudes. Cuando no es casual que los alumnos con mayores restricciones económicas sean precisamente los que tienen más dificultades para aprobar sus estudios, no es casual tampoco que esos alumnos sean desertores, reflejen un alto grado de ausentismo, no traigan los libros ni las tareas que los maestros les dejan; no es casual tampoco su falta de interés o de participación en clase. Sin embargo, resulta sospechosa la ausencia de estos análisis en los planteamientos de la evaluación.

Plantear la evaluación como una actividad social implica, por tanto reconocer las determinaciones sociales que la afectan. De esta manera, sería necesario pensar ¿qué papel desempeña la evaluación en un país capitalista dependiente?, ¿hasta dónde los sistemas de evaluación de estos países funcionan como una forma de legislación de la imposibilidad del

¹⁴ Véase M. Antebi y C. Carranza 1975, y S. Barco 1975.

¹⁵ A. Díaz Barriga 1978

¹⁶ L. Seve et. al 1978.

Estado para dar escuela a todos?¹⁷ No se puede negar el papel de control social que se ejerce con la evaluación, pues con ella se legitiman ciertas desigualdades sociales, ya que los más favorecidos en lo económico resultan escolarmente favorecidos; así, no podemos negar que los actuales sistemas de exámenes cumplen con esta función de control.

La evaluación, por lo tanto, no se puede analizar tan sólo como una propuesta técnica. Los textos de evaluación refuerzan habitualmente la idea de que esta actividad es científica, es técnica y tiene una neutralidad valorativa, definida por su pretensión de objetividad. De hecho, en la actualidad sería forzoso estudiar lo que los textos de evaluación no dicen sobre la evaluación, o sea, su silencio sobre la carga social implícita en ella. Con base en estos elementos de análisis es pertinente precisar que, tal como se encuentra actualmente, el discurso de la evaluación carece de una construcción teórica sobre su objeto de estudio, y se queda únicamente en un nivel de descripción y acción empiristas.

Segunda tesis: el discurso actual de la evaluación se fundamenta en la teoría de la medición, lo que impide el desarrollo de una teoría de la evaluación.

Esta proposición está configurada por dos elementos: (a) el discurso actual de la evaluación se fundamenta en la teoría de la medición; en la medición se encuentra la propuesta central para realizar la evaluación, por lo que este discurso manifiesta claras vinculaciones con el conductismo y con la concepción de ciencia del positivismo, y (b) la incorporación de la teoría de la medición en la evaluación impide que se desarrolle una teoría de la evaluación, en tanto que este discurso no abarca la construcción de su objeto de estudio, esto es, el estado del proceso del aprendizaje de un sujeto.

Habría que reconocer que en la evolución en la psicología a fines del siglo pasado y principios de éste, se ha operado un cambio en su objeto de estudio: el conductismo surgió y postuló como objeto la conducta observable: esto implica hacer un recorte sobre lo que era el objeto de estudio de la psicología —el hombre—, el cual es desgajado y dejado en un nivel de pseudoconcreción; en consecuencia, no hay en la psicología conductista una explicación teórica sobre los procesos psíquicos del hombre. De hecho hay una eliminación de aspectos sustantivos para la comprensión de estos procesos, como la categoría de lo inconsciente, porque se los considera transempíricos. En tanto la evaluación se fundamenta en la psicología conductista y, por lo tanto, en el empirismo, conlleva el mismo lastre frente a la ausencia de sus explicaciones teóricas. Es aquí donde surge la teoría de la medición como un elemento para «hacer como que se investiga», cuando en realidad lo único que se hace son mediciones empiristas.

De esta manera, el conductismo y la psicología fiscalista¹⁸ aplican para su «investigación» el método experimental. Fue precisamente con éste que apareció la teoría

¹⁷ Este papel resulta claro cuando examinamos algunas estadísticas de solicitudes de ingreso a la enseñanza media superior y superior, y las posibilidades reales de estas instituciones de admitir candidatos: las cifras de los lugares que se ofrecen son muy inferiores a las de las plazas solicitadas.

¹⁸ «en estas páginas, nos proponemos fundamentar la tesis de que toda proposición de psicología puede formularse en lenguaje físico porque todas las proposiciones de psicología describen acontecimientos físicos, a saber, la conducta física de los humanos y otros animales. Ésta es una tesis parcial de la tesis general del fiscalismo, conforme a la cual el lenguaje fiscalista es un lenguaje universal; esto es, un lenguaje al que puede

de la medición como un elemento que permitía el tratamiento estadístico de ciertos datos, pero, como establece Gloria Benedito,¹⁹ la psicología conductista toma la teoría de la medición de manera de formal, en tanto que no recurre a ella partiendo de las matemáticas, sino del método experimental que aparece en la fisiología pavloviana, que a su vez la toma de la física; esta última es la única ciencia que la toma directamente de las matemáticas. Así, la inclusión de la teoría de la medición en el discurso de la evaluación lleva el lastre de una serie de mediaciones previas que permiten explicar su deformación o, por lo menos, su simplificación. A la vez, el empleo de la estadística genera un «efecto de científicidad» en los planteamientos de la evaluación.

Si hacemos una lectura de los diferentes manuales que abordan el tema de la evaluación, descubriremos cómo sus fundamentos se relacionan con la distinción que hacen los autores entre medir y evaluar. Esta distinción se trata con toda claridad en el texto de Thorndike y Hagen, *Tests y técnicas de medición en psicología y educación*, cuyo capítulo primero se dedica a la explicación del empleo de la teoría de la medición. Así, en las definiciones de evaluación que dan diversos autores, aparece el problema de la medición; por ejemplo, para Mager²⁰ la evaluación «es el acto de comparar una medida con un estándar y emitir un juicio basado en la comparación», para Lafourcade²¹ «es una interpretación, de una medida (o medidas), en relación con una norma establecida».

Más aún, se puede hablar de una tendencia a utilizar el término medición como sinónimo de evaluación. Este empleo lo podemos observar en el texto ya citado de Thorndike y Hagen, y en el de Mager; también, con toda claridad, aparece en el artículo de Miel, «Tendencias actuales en la evaluación del rendimiento escolar», quien inicia el trabajo con la siguiente frase: «si se revisa la literatura concerniente a la medición de los últimos diez años...»²² estos elementos reflejan el grado de extensión y aceptación de esta tendencia.

Por ello se puede afirmar que la teoría de la medición se toma como referente único y sostén del discurso relacionado con la evaluación y que no se ha desarrollado en la actualidad una teoría de la evaluación. A la vez, la inclusión de la medición implica la aceptación de la posibilidad de medir el aprendizaje, por lo cual se recorta la misma noción de aprendizaje para concebirlo únicamente como los cambios de conducta más o menos permanentes en el alumno. La cuantificación del aprendizaje lo reduce a lo superficial y aparente, en detrimento del proceso mismo de aprender, de manera que esta «cuantificación en pos del prestigio de lo incuantificable es, si acaso, comparable al intento de Leibniz en demostrar matemáticamente la existencia de Dios»²³

traducirse cualquier proposición. Los partidarios del conductismo llegaron a su proposición a través del interés tenido por la psicología animal; en este terreno, donde no hay enunciados que observar, sino sólo la conducta sin palabras, es más fácil llegar al enfoque correcto. Este enfoque conduce a la interpretación correcta de los enunciados de los sujetos experimentales humanos» R. carnal 1978, pp. 169-204

¹⁹ B. Benedito 1975.

²⁰ R. Pager 1975, p. 20

²¹ P. Lafourcade 1973, p. 17.

²² E. Miel 1979, pp. 171-194

²³ G. Devereux 1977, p. 30. Un poco más adelante el auto señala: «El último punto es la ficción, aceptada sólo por los no matemáticos, de que un enunciado matemático o estadístico necesariamente tiene sentido. Es simplemente una versión moderna de la opinión de que cualquier cosa dicha en latín tiene sentido»

Una vez aceptada la medición en el discurso de la evaluación, la «tecnología» sugerida para realizar esta actividad reproduce la misma lógica del conductismo. Por ello, Thorndike y Hagen exponen el sencillo procedimiento que hay que seguir en la evaluación.

Medir consta de tres pasos: (a) señalar y definir la cualidad a medir; (b) determinar el conjunto de oraciones en virtud de las cuales el atributo puede manifestarse y hacerse perceptible, y (c) establecer el conjunto de procedimientos o de definiciones para traducir las observaciones a enunciados de procedimientos de grado o cantidad.²⁴

Es necesario hacer notar la semejanza de este planteamiento con las formulaciones del método experimental, fundamentalmente cuando se aplica a la conducta humana. Así, encontramos que los autores que estudian este método proponen que las hipótesis formuladas para una investigación «establezcan una relación entre dos o más variables que sean mensurables o potencialmente mensurables».²⁵

Esta lógica se expresa claramente cuando se trata de formular un proyecto de evaluación. Así, Weiss dice:

La pregunta evaluativa parece ser muy sencilla, lo único que tiene que hacer el evaluador es: (1) buscar las metas del programa; (2) traducir las metas a indicadores mensurables de la realización del objetivo; (3) reunir datos de los indicadores, y (4) comparar datos con criterios de realización de las metas.²⁶

De esta forma, la cuantificación del dato se convierte en la técnica apropiada para la verificación. Este procedimiento, vinculado directamente al positivismo, reduce el problema del conocimiento al ámbito fenomenológico. En términos de Kosik, este dato no es más que una pseudoconcreción; sin embargo, tal como aparece en este planteamiento, no se requiere llegar a la explicación última de los fenómenos, o sea, a su esencia, en tanto que lo que importa es la eficacia²⁷ de las técnicas empleadas.

Quisiéramos también exponer cómo funciona el modelo propuesto en la práctica de la evaluación por Thorndike, a partir de los tres pasos antes definidos. Si consideramos un curso de filosofía, en el primer paso podríamos afirmar que el estudiante tiene que aprender filosofía moderna, con lo cual podríamos empezar a desarrollar el segundo paso, que indica que se deben «determinar el conjunto de operaciones que permitan que este aprendizaje se manifieste». Así, aprender filosofía moderna queda reducido a un conjunto de desempeños comportamentales o bien de objetivos de unidad; por ejemplo, que el estudiante indique el año en que nació y murió Hegel, que mencione tres de sus principales obras, describa dos tesis fundamentales de su pensamiento, etc. Una vez cubierto este paso, se efectúa el tercero, el cual tiene que ver prácticamente con la cuantificación de los datos, que en este ejemplo se traduce en la elaboración y aplicación de un conjunto de ítems.

Es necesario hacer notar el salto epistemológico que se produce en los pasos 1 y 2. Esto es, que el atributo sólo lo conocemos por la definición operacional del mismo. Sin embargo, por las características de esta definición operacional que obligan a una

²⁴ R. Thorndike y e. Hagen 1970, p. 17

²⁵ E. Kerlinger 1979, p. 18.

²⁶ C. Weiss 1975, p. 39.

²⁷ «Una cosa es atacar ciertos puntos de vista de Skinner, y otra muy distinta negar la eficacia probada del modelo operante en la investigación, en la enseñanza» R. Bavés 1970, p. 29.

formulación «observable», el atributo sólo es «conocido» por sus apariencias. Con este ejemplo queremos mostrar cómo difícilmente se puede aceptar que un estudiante haya aprendido filosofía moderna a partir de los indicadores que se derivarían de una serie de desempeños u objetivos de unidad. Esta dificultad epistemológica es grave en tanto que condena a ese modelo a moverse únicamente en la pseudoconcreción o apariencia de un fenómeno. Por ello, cabe afirmar que la inclusión de la teoría de la medición en la evaluación le confiere a ésta sólo un efecto de cientificidad.²⁸

Lo anterior explicaría por qué existen tan pocas diferencias en los textos de evaluación. Las propuestas que en ello se hacen no son sustancialmente diferentes, puesto que en tales textos, que fundamentan la evaluación en la medición, se dan a conocer técnicas de interpretación y de tratamiento de la información, como la obtención de la media aritmética, el modo, la desviación estándar, etc. en la práctica, la propuesta de evaluación del aprendizaje se reduce a la confección, aplicación e interpretación de datos recabados por medio de exámenes, los que se aplicarían al alumno al finalizar la enseñanza.

Han alcanzado gran difusión las llamadas pruebas «objetivas», que pueden ser elaboradas e interpretadas mediante sistemas mecanizados, con lo cual se minimiza tanto el proceso mismo de la evaluación del aprendizaje como la noción de aprendizaje y la de docencia. Aquí surge el problema de la «objetividad» en la evaluación, objetividad que en los textos se reduce únicamente al tercer paso de Thorndike, esto es, frente a la dificultad epistemológica que existe entre el atributo y su definición operacional, los textos plantean la posibilidad de la objetividad, reducida únicamente al momento de la cuantificación de los datos.

Por ello, este planteamiento en relación con la objetividad es un problema que, en cuanto a paradigma, produce del positivismo y que, en el contexto en el que se maneja, es únicamente «objetivismo» que impide comprender y explicar el proceso del aprendizaje de un grupo o de un individuo. En estos planteamientos se omite la relación sujeto-objeto en el acto de conocimiento y las mediaciones que se dan en la misma acción de conocer, a la vez que se niega la riqueza que existe entre observador y observado en las ciencias del comportamiento, elemento que tiene que ver con la conceptualización freudiana de transferencia y contratransferencia. Así, es necesario reconocer que en el caso del comportamiento humano «nunca se observa un hecho que hubiera producido en ausencia. Por fortuna, los llamados trastornos y perturbaciones creados por la existencia del observador, son piedras angulares de una verdadera ciencia del comportamiento y no contratiempos deplorables.»²⁹

Al desconocer estos planteamientos, el actual discurso de la evaluación cae en un «objetivismo» parcial, en tanto que lo que se objetiva es el momento de asignar una expresión numérica al alumno, de acuerdo con la cantidad de preguntas contestadas correctamente; pero no existe tal objetividad ni en el muestreo de los aprendizajes que se deben lograr, ni en la misma elaboración de las preguntas.

No hay que perder de vista que estas pruebas objetivas difícilmente sirven para medir aprendizajes que no sean, en última instancia, memorísticos; por lo tanto, no

²⁸ «Por física entendemos aquella ciencia caracterizada por su procedimiento para la formación de conceptos: reduce todo concepto a relaciones de magnitud, esto es, a una sistemática atribución de números a puntos espacio-temporales; entendida así la física, podremos expresar nuestra tesis —tesis parcial de fisicalismo—, del modo siguiente: La psicología es una rama de la física» R. Carnal 1978

²⁹ G. Devereux 1977, p. 31.

fomentan el desarrollo de las capacidad crítica y creativa, la resolución de problemas, el manejo de relaciones abstractas, etcétera.

Si bien algunos autores presentan ejemplos de reactivos que no reflejan únicamente memorización del material, por el número reducido de estos casos, cabe afirmar que constituyen la excepción que confirma la regla. Bastaría con observar los reactivos de los exámenes departamentales para llegar a esta conclusión. Con esto cobra significado la afirmación hecha por Bloom durante el «Simposium sobre Currículum Universitario» expresa:

Las pruebas hechas por profesores, las estandarizadas, son en gran medida, pruebas de conocimientos memorizados. Después de más de 20 años del uso de la taxonomía de objetivos educacionales en el adiestramiento de profesores, más del 95% de las preguntas contenidas en las pruebas que los estudiantes deben contestar se refieren a poco más que la mera información.³⁰

Así, el problema de evaluación se ha restringido al empleo de este tipo de pruebas, a la sofisticación del proceso, con su paralela mecanización, y no se ha llevado a la comprensión del proceso de aprendizaje, tanto individual como grupal. También se ha generalizado la idea de que con el empleo de la estadística y, en última instancia, de la cuantificación, la evaluación adquiere —como lo hizo en su tiempo la psicología empirista— el estatus de actividad científica.

Es necesario subrayar este error, ya que en la cuantificación de fenómenos del aprendizaje «no se respetan las propiedades de los números».³¹ ¿Cuál es el significado de un 6 en anatomía o en Cálculo?, ¿cuál es la diferencia entre un 8 y un 10?, ¿qué propiedad del número se está empleando en esta designación?, ¿qué significa tener 64 aciertos en 80 reactivos? No se pueden responder estas preguntas argumentando que ahora, en algunas instituciones —particularmente en la UNAM—, se emplean factores simbólicos: MB, B, S, NA, para designar los niveles de aprovechamiento, puesto que estos factores representan un número (no una propiedad del número), y cuando se pretende manejarlos, por ejemplo en la obtención de promedios, se convierten aritméticamente en el número que representan. Esto revela que el empleo de fórmulas estadísticas para determinar el grado de «validez» y «confiabilidad» de las pruebas llamadas «objetivas» ha cobrado amplia difusión.

Frente a este hecho haremos algunas precisiones. En primer lugar, el estudio estadístico de tales pruebas es un estudio centrado en el instrumento (la prueba), con lo cual el docente pierde de vista la totalidad del proceso, tanto del aprendizaje como de la evaluación escolar. La importancia que se da al análisis del instrumento propicia que, una vez perfeccionado dicho instrumento, se aplique indistintamente —en supuestas condiciones de igualdad— en un curso de Historia de México o en uno de Anatomía. Lo único que el docente tendría que hacer, desde esta perspectiva, es solicitar el instrumento —la prueba— elaborado para el curso, de la misma manera en que se pide un libro para esa materia.

Los exámenes departamentales y los exámenes nacionales de medición de la calidad cumplen un a función similar cuando, ajenos a las condiciones particulares de un proceso grupal particular y del proceso del aprendizaje que se haya generado, presentan las mismas

³⁰ B. Bloom 1978.

³¹ N. Braunstein *et al* 1975.

preguntas a los grupos que cursan la misma materia. Esto explicaría por qué tienen que recurrir únicamente a preguntas que sólo requieren la memorización de una información, y permitiría pensar que más que una ayuda para el proceso del aprendizaje cumplen una función de control sobre los docentes. Esto se inserta en una concepción que reduce las funciones docentes a acciones netamente mecánicas, porque el maestro ya no tendría que averiguar los logros del alumno, ni qué situaciones se facilitaron o se dificultaron, cómo se dio el proceso grupal, etc. Esto implica que, así como se han elaborado «máquinas de enseñar» que de manera mecánica presentan cierta información al estudiante, se aplican los mismos principios lógicos para «medir» el grado de retención de esta información; esto, de manera implícita, indica que las funciones docentes son puramente mecánicas y estáticas. Con ello se rompe la dialéctica propia de una situación de enseñanza-aprendizaje, que coloca al grupo escolar humano en condiciones de realizar múltiples actividades que configuran al aprendizaje como un conjunto de acciones técnicas e irrepetibles en un proceso dinámico.

El planteamiento anterior se vincula directamente con la lógica de la pedagogía pragmática de este siglo, que minimiza el papel y la función del maestro al restringirlo únicamente a un «instrumento» que supervisa, guía, conduce lo que los planificadores han establecido. En cierta forma, ésta es una lógica para que el maestro no piense.³² Las principales consecuencias de esta visión estática están dadas por una consideración mecánica de la docencia y del aprendizaje, y por una visión mecánica del hombre, al que se considera como un sistema de almacenamiento y emisión de información. El poder de discriminación de un instrumento «objetivo» está determinado por diferencias de puntaje que alcanzan aquellos que contestaron «mejor» la prueba —los cuales obtuvieron calificaciones más altas— y los que lograron las puntuaciones más bajas.

Aparece aquí el problema de la *evaluación realizada con referencia a una norma*, donde lo importante es efectuar una serie de comparaciones entre los individuos y no captar los resultados obtenidos por cada estudiante en el curso del aprendizaje. El problema se repite en sus características fundamentales: lo importante es validar el instrumento para que éste se pueda generalizar. Es tan importante este dato, que no importa cuáles hayan sido las causas que incidieron en el éxito de quienes mostraron alto rendimiento, o qué determinó el fracaso de los que lograron la calificación más baja. En última instancia, es necesario replantear el problema de la evaluación escolar, de suerte que las opciones que se ofrezcan para resolverlo no se reduzcan al estudio y al análisis de los instrumentos de evaluación. Hemos visto que si en este replanteamiento no se cuestionan los fundamentos psicológicos y científicos de la propuesta de evaluación, las opticores que se den resultarán redundantes. En última instancia, este esclarecimiento teórico debe repercutir en la realización de experiencias diferentes en las diversas situaciones de aprendizaje. Es necesario puntualizar la noción de aprendizaje, de conducta y de personalidad; definir la dinámica del aprendizaje escolar, y el significado y el valor del grupo en el mismo; encontrar fundamentos

³² Esta idea es muy clara en el trabajo de Snyders, en el que establece que en la escuela tradicional se asigna una función profesional al docente y, por lo tanto, éste es el sujeto responsable de un conjunto de decisiones profesionales sobre el proceso escolar (selección de contenidos, tanto informativos como valorativos y de procedimiento), elaboración de la propuesta metodológica de trabajo de acuerdo con las condiciones de cada grupo escolar. Por el contrario, con los planteamientos derivados de la pedagogía pragmática estadounidense, el docente queda restringido a un «operario en línea de producción escolar», que debe aplicar lo que fija el Ministerio de educación, los especialistas en programas y los expertos en elaboración de pruebas de medición de la calidad, con las que finalmente se medirá el desempeño de sus alumnos y su propio desempeño.

epistemológicos para el objeto de la evaluación, que permitan una reconstrucción de su discurso que supere las propuestas empírico-analíticas, e incorpore el estudio en la problemática de la *comprensión*³³ de las causas del aprendizaje, de las condiciones internas y externas que lo hacen posible, y del proceso grupal en el que tiene lugar.

De este modo, la evaluación educativa no puede ser realizada únicamente por expertos, departamentos de evaluación, etc., sino que deben ser los docentes y los alumnos quienes participen, de manera privilegiada, en ella. El objeto de la evaluación se traduce así en la indagación sobre el proceso del aprendizaje de un sujeto o de un grupo, indagación que permite detectar las características de este proceso y buscar una explicación a las mismas, y que rebase la parcialidad de atender sólo algunos resultados del aprendizaje.

ALGUNAS DERIVACIONES DE ESTAS TESIS EN LA DOCENCIA

En esta segunda parte intentaremos dar un conjunto de respuestas prácticas que puedan ser utilizadas por los docentes en cada una de sus situaciones escolares, ya que finalmente son ellos los que tienen que resolver los problemas que se presentan en cada grupo escolar. Estas respuestas constituyen derivaciones de la concepción que previamente sustentamos, y por ello debemos considerarlas como propuestas y sugerencias que se mueven en un plano tentativo, ya que ni agotan las posibilidades que se pueden desprender de nuestra concepción teórico-política de la evaluación, ni funcionan como criterio de verdad. Más bien pretendemos señalar el estado actual de nuestras reflexiones que no incluyen propuestas cerradas ni acabadas, sino ideas abiertas a la posibilidad de desarrollos posteriores.

Primera: establecer una distinción práctica entre evaluación y acreditación

Consideramos que al establecer esta distinción, desarrollamos un elemento útil para circunscribir los problemas de certificación, aprobación y asignación de calificaciones al ámbito institucional al cual pertenecen, de suerte que no afecten ciertos desarrollos teóricos y prácticos respecto a la evaluación, sobre todo en los elementos que aportan para la comprensión del proceso educativo, en particular del proceso del aprendizaje escolar. Esta distinción nos permite determinar el papel institucional que tienen los reglamentos de exámenes, como elementos normativos que permiten el desarrollo de un sistema escolar y que no forzosamente coinciden con ciertos planteamientos teóricos sobre la concepción de la evaluación. La acreditación se relaciona con la necesidad institucional de certificar los conocimientos; está referida a ciertos resultados de aprendizaje que a veces están contemplados en los criterios de acreditación de un curso, pero que no dejan de ser cortes artificiales en el proceso de aprendizaje de una persona.

En ocasiones se llega a afirmar que los alumnos lograron algunos aprendizajes; esta forma coloquial puede originar que se pierda de vista que el aprendizaje es un proceso y no una cosa que se puede tener. En el fondo de esta problemática continúa presente una perspectiva conductista, en la que se concibe el aprendizaje como producto, en contraposición a otras explicaciones que resaltan el proceso sobre el producto

³³ El término debe entenderse en el marco de la moderna teoría del conocimiento, fundada en el debate alemán del siglo pasado. No se trata de ubicar una relación causa-efecto, sino de reconocer la complejidad de factores que afectan el proceso de aprender.

Estos *cortes artificiales* en el proceso de aprendizaje de un estudiante permiten derivar las calificaciones, lo cual se relaciona más con una necesidad institucional de respetar ciertos tiempos (semestrales) del calendario escolar, que con una necesidad el proceso de aprendizaje del estudiante.

En capítulos anteriores hemos sostenido que los criterios de acreditación, necesarios desde un punto de vista institucional, son sólo cortes artificiales en el proceso de aprendizaje de un estudiante. Dichos criterios establecen algunos productos o resultados³⁴ que el estudiante deberá mostrar en determinado tiempo; sin embargo, si aceptamos que el aprendizaje es un proceso en el que el estudiante confronta su esquema referencial con cierta información que moviliza el mismo esquema, y que a través de una elaboración de la información produce una nueva organización de éste, ciertamente los resultados de aprendizaje de un estudiante sólo reflejan un momento de este proceso.

Concebir al aprendizaje como producto es una manera de desconocer que toda persona necesita combinar una etapa de recepción de nueva información con otra etapa —lenta e imperceptible— de elaborarla, con el fin de construir una síntesis propia. Es la diferencia entre repetir sin entender (repetir «como perico») y producir un discurso propio, el cual, a fin de cuentas, se basa en las ideas que cada sujeto ha ido estructurando. La estructuración de nuevas ideas se da a partir de la mediación del esquema referencial de cada sujeto, esto es, del conjunto de conocimientos, sentimientos y afectos con los que piensa y actúa, esquema que ha sido organizado a lo largo de la vida del propio sujeto y que, al entrar en contacto con una nueva información, es susceptible de ser movilizado y reestructurado. «Lo importante —señala Bleger— no es sólo el esquema referencial consciente, sino todos sus componentes inconscientes o disociados que entran en juego y que, desconocidos, distorsionan o bloquean el aprendizaje. En una buena proporción, el esquema referencial es el *a priori* irracional del conocimiento racional y de la tarea científica.»³⁵

Es importante considerar que los criterios de acreditación deben establecer pruebas de aprendizaje que permitan la integración del conocimiento, para evitar el desarrollo de comportamientos que tiendan hacia la fragmentación. Obviamente los exámenes nacionales de calidad profesional tienden a privilegiar lo fragmentario en detrimento de tal integración. Esas pruebas, construidas con reactivos de «opción múltiple», parten de la idea de que es conveniente hacer un muestreo de la información que contiene un programa o una etapa de formación. Esta situación impide promover y, en su caso, detectar los procesos de aprendizaje del estudiante, ya que fomenta la memorización y la mecanización tanto del proceso del aprendizaje del alumno, como de las actividades de aprendizaje que propone el docente. Estos planteamientos, en fin, dejan de lado una serie de aprendizajes reales que se vinculan a las capacidades específicamente humanas, como el juicio crítico y la capacidad de establecer relaciones. Por ello, aun cuando una persona pueda contestar un cuestionario en el que se le pide que recuerde cierta información, de sus respuestas no se obtienen datos sobre su proceso de aprendizaje, ni sobre las dificultades que enfrenta. Cabe reiterar que sólo con la especificación de lineamientos de acreditación integrados se pueden promover

³⁴ Utilizamos la noción de mundo externo del modelo de áreas de la conducta de Bleger, que la refiere al área donde aparecen —se objetivan— los productos de la conducta humana, sin negar la unidad que guardan con las otras dos áreas: cuerpo y mundo externo. Véase J. Bleger 1976^a, pp. 55-75.

³⁵ J: Bleger 1976b, pp. 70-71 (las cursivas son mías)

algunas manifestaciones de aprendizaje que el alumno muestre al término del curso, sin confundirlas con otros procesos generados en los mismos estudiantes.

Las actividades que se desprenden de los criterios de acreditación no llevan necesariamente a un trabajo final o a un examen de término de curso. De hecho, estos criterios pueden afectar la organización y planificación de actividades de aprendizaje a lo largo del curso, aunque no constituyen el único referente para esta planificación. En todo caso, la acreditación de un curso responde a un análisis que inicialmente realiza el docente, en academia o de forma individual, sobre los elementos más significativos que tiene el contenido del curso para el estudiante, así como la adopción de un concepto de aprendizaje por parte del docente. Estos elementos lo ayudan a establecer la evidencia o conjunto de evidencias que los alumnos mostrarán al término del curso, conjunto que puede ir construyendo a lo largo del trabajo cotidiano, desde el primer día de clases, y para lo cual es importante que el programa del curso por desarrollar contenga de manera muy clara y detallada los criterios de acreditación. Ya hemos afirmado que éstos constituyen una forma de contrato entre docentes y estudiantes. En el caso de la educación superior o de la educación para el trabajo, es conveniente considerar algunos elementos referidos a práctica profesional como orientadores centrales en la definición de tales criterios.

La diferenciación entre acreditación y evaluación nos ayuda a no mezclar elementos formativos con elementos de aprendizaje en el acto de acreditación. Algunos aspectos formativos como limpieza, puntualidad, orden y dedicación, necesitan ser fomentados en procedimientos de evaluación del trabajo escolar, pero no deben formar parte de la acreditación de un curso. Esto significa reconocer que es posible que un alumno pueda estar muy bien acreditado pero muy mal evaluado, porque de hecho puede cumplir con los requisitos para acreditar un curso, referidos a las precisiones establecidas en los criterios de acreditación y mostrar algunos de vicios o carencias en su proceso de aprendizaje³⁶ Esta disociación entre acreditación y evaluación plantea otro problema: ¿dónde se sitúan aquellos aspectos que no están presente en la acreditación y que son los más importantes del proceso de aprendizaje? La respuesta la hemos esbozado a lo largo de nuestra exposición, ya que consideramos que la *evaluación* permitiría actuar conforme a estos aspectos, de manera que el objeto de estudio de la evaluación fuera la comprensión del proceso de aprendizaje en el grupo y en los individuos.

Lo anterior implica que un alumno que obtiene la calificación más alta —aun en los exámenes del Centro Nacional de Evaluación mexicano (CENEVAL), que cada vez es más proclive, dar a conocer sus nombres a la prensa— no es necesariamente el que desarrolló más su proceso de aprendizaje ni sus capacidades analítico-creativas, no el que más colaboró en el proceso de aprendizaje ni sus capacidades analítico-creativas, ni el que más colaboró en el proceso de aprendizaje grupal. Existen muchas formas de colaborar en el aprendizaje en grupo: la realización consiste y amplia de lecturas, la propuesta de enfoques de discusión o de interpretación de inmaterial, las interrogantes que plantea cada integrante de un grupo escolar son, entre otros, factores de enriquecimiento y potenciación del grupo; sin embargo, las actuales políticas educativas en México tienden a desconocerlos. A veces

³⁶ Villarroel establece: «si usted recibe un trabajo sucio del alumno, esto en nada invalida el conjunto de operaciones aritméticas que realizó para llegar a un resultado, ni los razonamientos implícitos en ellos, un trabajo sucio puede causar una mala impresión estética.» Véase C. Villarroel 1974, p.122. en mi experiencia profesional he tenido varios casos de estudiantes que han cumplido muy bien con los requisitos que se establecen en los criterios de acreditación y, sin embargo, reconocen que a lo largo del curso no lograron avanzar en su proceso de aprendizaje. Véase A. Díaz Barriga 1994^a, cap. 6

los alumnos que obtienen mejores calificaciones son los que menos colaboran o los que menos esfuerzos realizan. Cuando se considera que el objeto de la evaluación es la comprensión del proceso de aprendizaje, se busca responder las siguientes interrogantes: ¿por qué este proceso se dio de determinada manera?, ¿que aspectos del curso favorecieron o dificultaron el proceso de aprendizaje?, ¿hasta qué punto el estudiante se percibe a sí mismo en su proceso de aprender?

Las prácticas escolares que actualmente se realizan fomentan la individualización de este proceso, en el que el estudiante se preocupa más por obtener una «buena» calificación, que por conocerse, observarse y reflexionar sobre su aprendizaje. Estas prácticas tienden a deformar la percepción que del *yo* hace el sujeto, propiciando que no sólo intente engañar a los demás en relación con su aprendizaje sino que, aun por mecanismos no conscientes para el mismo sujeto, se engañe en cuanto a su mismo aprendizaje, creyendo que ya aprendió. Si consideramos que el aprendizaje es un proceso del sujeto, expresiones como «ya aprendió», «ya logró los objetivos del curso», carecen de sentido y significado, puesto que toda información que el sujeto recibe puede ser continuamente reelaborada por él en la interrelación que se establece entre información y «esquema referencial» del sujeto, en un ajuste continuo del mismo. La función de la evaluación es indagar los ajustes que se van suscitando en el mismo esquema y, fundamentalmente, las causas de los mismos. La evaluación, vista desde esta perspectiva, es una interrelación entre una evaluación individual y otra grupal; es un proceso que permite a cada integrante de un grupo escolar reflexionar sobre su aprendizaje para confrontarlo con el seguido por los demás miembros del grupo; a la vez, permite desarrollar una percepción desde el punto de vista del grupo sobre el desarrollo de cada uno de sus miembros. Hay menos engaño entre pares que el que deriva entre una calificación otorgada por un maestro o un sistema de exámenes y los estudiantes vistos en una especie de anonimato. De esta manera, la evaluación tendería a propiciar, en cada alumno, una percepción de su propio proceso de aprendizaje.

En ocasiones, cuando se pide al estudiante que realice un trabajo para acreditar el curso, se considera que este trabajo es algo acabado y no la muestra de un momento particular en la actividad de elaboración de la información, momento que permite tener evidencias de ciertas aproximaciones a un objeto de estudio, pero que a la vez conserva ciertas imprecisiones e incongruencias que es necesario seguir elaborando. Así, el aprendizaje no es algo acabado. En este sentido, la evaluación permitiría que el estudiante y el grupo escolar reflexionasen sobre lo que se alcanzó a integrar en un curso y lo que no se pudo integrar, así como sobre las causas de ello; todo esto permitiría tomar decisiones sobre qué hacer en un futuro.

Segunda; el examen es un instrumento inadecuado para realizar la evaluación y no es el más apropiado para obtener las evidencias para la acreditación.

Esta afirmación tiene que ver con la manera en que se efectúan habitualmente las pruebas en los diferentes cursos, en los cuales el alumno se enfrenta a una serie de preguntas a las que tiene que dar respuesta en un tiempo determinado, dentro del salón de clases y bajo la vigilancia de un maestro. Paradójicamente, la literatura sobre la evaluación escolar, como ya lo explicamos, dedica buena parte de su espacio a la elaboración y manejo de resultados de las pruebas. Sin embargo, poco se ha reflexionado acerca de cómo la propia historia del

examen muestra que éste ha sido empleado como un instrumento de control social, de elección de elites; una mirada histórica sobre la práctica educativa del examen mostraría que este instrumento no se empleó inicialmente para certificar conocimientos, sino como parte de la metodología de enseñanza, cuya función era garantizar el aprendizaje, o bien como el momento de mostrar en público la madurez académica alcanzada.³⁷ Judges, en su historia sobre este instrumento, enuncia que sus orígenes se pueden encontrar en el siglo XII antes de nuestra era, cuando los mandarines chinos aplicaron, por primera vez en la historia de la humanidad, exámenes a miembros de otras castas para elegir a quienes habrían de formar parte de la burocracia. El autor considera que en la época victoriana se implantó el examen como estabilizador de los miembros de la jerarquía en el poder; en las universidades de Oxford y Cambridge, el examen de honores consistía en averiguar el grado de competencia que había logrado un candidato, o si era competente para recibir un título. Sin embargo, el Examination Statute de Oxford (1800) establecía, por acuerdo general, que sólo una minoría de aspirantes al grado o título pudiera disfrutar de este privilegio; los jesuitas extendieron la competencia al más alto grado de efectividad desarrollado hasta entonces, aunque según Foucault,³⁸ es La Salle quien establece la escuela como un aparato de examinación continua. Durkheim³⁹ describe cómo el examen de la universidad medieval del siglo XII se degenera en el siglo CSB, para convertirse en un ritual de ridiculización del estudiante. A mediados del siglo XIX, el examen adquiere la forma que conocemos, al vincularse una calificación que determina una promoción escolar. Para algunos autores este hecho significó que tanto alumnos como maestros perdieran el sentido de la educación, el placer de aprender y placer de enseñar.⁴⁰

A principios del siglo XX, Bidet, en sus investigaciones sobre la inteligencia en deficientes mentales, abrió de manera no consciente el camino para sorprendentes aplicaciones al campo de los exámenes; la construcción de tests de inteligencia paulatinamente fue dando pauta a la construcción de tests de rendimiento escolar. Aunque el debate pedagógico estadounidense de este siglo está marcado por los excesos y vicios generados en el sistema social y en el sistema pedagógico sobre estos instrumentos, los tests de rendimiento escolar se han aplicado cada vez en mayor número de escuelas. Se puede consultar la historia de la educación de este siglo en los Estados Unidos, para analizar el conjunto de vicios, errores y excesos que se cometieron a partir de los tests. Cremin⁴¹ muestra lo que denomina el paso sutil de la ingeniería a la pedagogía, afirma que los resultados de los tests se empleaban para justificar la segregación educativa. Bowles y Gintis⁴² analizan cómo se empleaban los tests para justificar la marginación social de prostitutas y reos negros; Kemmis⁴³ analiza cómo los psicólogos eran llamados a la corte para testificar en contra de deficientes mentales, de suerte que pudieran ser castrados para evitar su reproducción social. Así los tests, como forma de medir científicamente al ser humano, cumplieron la función de control social que el sistema les asignaba. Hoy, en los

³⁷ Una mirada histórica mucho más detallada se puede obtener en la antología: A. Díaz Barriga (comp.) 1993^a.

³⁸ M. Foucault 1978. Para el autor, cuando La Salle elabora la *Guía de las escuelas cristianas*, establece un aparato de examinación permanente en las escuelas. Todos los días los alumnos deberán presentar un examen y habrá días en que deban presentar dos.

³⁹ E. Durkheim 1981.

⁴⁰ F. Giner de los Ríos 1906.

⁴¹ L. Cremin 1961

⁴² S Bowles y H. Gintis 1978

⁴³ Kemmis 1983.

planteamientos de Apple,⁴⁴ forman parte del programa neoconservador de la nueva derecha estadounidense.⁴⁵

Esta crítica a los exámenes vistos como instrumentos de control social nos vincula a las tesis previamente planteadas en este capítulo, dado que no se puede dejar de considerar que el examen confiere más bien un certificado social que una garantía de competencia técnica. En muchos empleos se puede observar que un puesto de conducción se acompaña de otro de ejecución, en el que existe el domino técnico. Al puesto de conducción suele, corresponderle la existencia de un título o certificado escolar que ampare determinado conocimiento, con la remuneración y el prestigio asociados, mientras que para el otro puesto, que en muchas ocasiones muestran un dominio efectivo de los procedimientos profesionales exigidos, se exige una certificación escolar menor.⁴⁶ La práctica de los exámenes fomenta la confusión y ansiedad en los estudiantes, y en ocasiones el docente la ve como una oportunidad para «desquitarse» de la falta de atención e interés de éstos. Si bien el examen es una de las causas por las que los alumnos «estudian», a veces sólo repiten información que no comprenden y que pronto olvidarán, o se preparan para realizar alguna práctica fraudulenta que les permita aprobar y, de alguna manera, quitarse de «encima» un examen. En este sentido, se puede entender por qué Landsheere los llama «cuerpos extraños en la educación, al servicio de una pedagogía anticuada».⁴⁷

Existen algunas experiencias que permitirían replantear este tipo de prácticas, por ejemplo, la que describen Mirta Antebi y Cristina Carranza, en la que se incorpora el empleo de la bibliografía en el momento del examen, lo que significa que los estudiantes trabajen directamente —como en la vida real— con la información, y no se les pide que la reproduzcan sólo como un acto memorístico. Esta experiencia describe también la posibilidad de recurrir al trabajo en equipo como elemento que potencia el manejo de cierta información e incita a la discusión de un problema —lo que también se observa en el trabajo cotidiano de empresas de punta—; más aún, las instrucciones dadas a los alumnos establecen que, cuando no haya acuerdo, se indique en un informe el desacuerdo y los motivos que lo fundamentan.⁴⁸

En la práctica docente también hemos tenido oportunidad de realizar diversas experiencias en el proceso de examinación, de ellas podemos concluir lo siguiente: (a) la modificación del proceso de examinación y de acreditación es el resultado de una modificación estructural en la parte metodológica del curso: no se puede pedir una actitud participativa, analítica y creativa en el curso y concluir con un examen de devolución pasiva de conocimientos; (b) a mayor compromiso de docentes y alumnos —como forma de grupo escolar—, la mayor posibilidad de idear estrategias creativas para la acreditación de un curso. Varios de nuestros colegas docentes también realizan experiencias innovadoras en este ámbito. En todo caso, la generalización de exámenes de «medición de la calidad», que finalmente se reducen a una prueba objetiva lejos de contribuir a la innovación de estas formas educativas, promoverá el achatamiento de las mismas.

Si el valor de un curso escolar o de un plantel educativo se va a determinar sólo por los resultados que obtengan los alumnos en estos exámenes nacionales, se dará pie a

⁴⁴ M. Apple 1993.

⁴⁵ Véase A. Judges 1971, pp. 33-57. En este capítulo, Judges hace una descripción más detallada que la que hemos resumido en este párrafo.

⁴⁶ G. Landsheere 1972, p. 33.

⁴⁷ *Ibid.*, p. 11.

⁴⁸ M. Antebi y C. Carranza 1975, pp. 59-90

generalizar una serie de vicios, tanto en el plano pedagógico como vinculados a la sobrevivencia. En el plano pedagógico podemos señalar dos: (a) entrenar a los alumnos cada semana o cada mes para que resuelvan pruebas similares, esto es, con preguntas parecidas o idénticas a las que se aplican en el examen nacional.⁴⁹ Y en lo cual se usarán los manuales que se empiecen a editar como «guías para preparar el examen de matemáticas del sexto grado»,⁵⁰ y (b) se propondrán cursos para «preparar» a los alumnos para resolver estos exámenes, en los que se harán exámenes con preguntas similares, tomadas de los múltiples manuales existentes; también se pueden dar elementos técnicos de construcción de una prueba objetiva, ya que éstos ayudan a resolverla con mayor eficiencia.⁵¹

En el plano estratégico han aparecido elementos que indican lo desastroso que pueden ser este tipo de exámenes en nuestro medio. En algunas escuelas chilenas se «expulsa» al niño que obtiene rendimiento bajo en el «examen de medición de la calidad», porque baja el promedio de la escuela; en otras se hace que los niños que obtienen menor puntaje no asistan a la escuela en la semana en la que se aplicarán tales exámenes. Finalmente, el sentido social y público de la educación queda pervertido en ambas prácticas.⁵²

Por estas razones, entre muchas más, lamentamos que en esa época de examinación estatal no se contemple ni el sentido social de la educación, ni la posibilidad de prácticas pedagógicas que atiendan más a los procesos educacionales que a los resultados; lamentamos que el efecto de los exámenes nacionales sea «achatar» el universo pedagógico para tener gráficas nacionales más estables. Seguimos preguntándonos, ¿hasta qué punto pueden participar los alumnos en la identificación de los problemas por resolver?, ¿hasta dónde los alumnos estarían en condiciones de revisar el trabajo de otro equipo que no sea el suyo, o por lo menos de opinar sobre él?, ¿hasta dónde lo fundamental de esta experiencia se podría realizar fuera de una situación clásica de examen, por ejemplo, fuera del aula, dando cierto tiempo para realizar un trabajo escrito, etcétera?

Si pensamos que no es con un examen como los alumnos mejoran su proceso de aprendizaje, sino que es necesario modificar las concepciones que los sistemas social y educativo, así como los docentes y los alumnos, tienen sobre el aprendizaje, entonces se imponen modificaciones sustanciales a las prácticas educativas que los docentes realizan en

⁴⁹ Recuerdo a un grupo universitario que fue injustamente sancionado por haber reproducido un examen objetivo. No se trata de que se hayan «robado» o «fotocopiado» el examen, sino de que se organizaran para, a la salida del examen de otros compañeros, preguntarles sobre qué temas habían sido examinados y de qué preguntas se acordaban; al cabo de varios meses tenían elementos adecuados para «entrenarse» en el examen. No habían hecho «trampa», en mi opinión, habían mostrado ingenio.

⁵⁰ Una visita a las librerías de Santiago de Chile permite vislumbrar lo que será esta industria editorial en nuestros países. En esas librerías hay estantes con guías —manuales de preguntas y respuestas construidas en la perspectiva de una prueba objetiva— para que los alumnos preparen esos exámenes. También han proliferado múltiples «academias» que ofrecen entrenar al alumno en la resolución de estas pruebas.

⁵¹ Por ejemplo: el distractor no puede estar en el mismo lugar en dos preguntas seguidas, entonces se sugiere al alumno que conteste rápidamente todas las preguntas de las que está seguro, luego que siga con aquellas en las que tiene alguna duda. Si en la pregunta anterior la respuesta correcta fue «b» y en la pregunta siguiente la correcta es «d», el alumno debe buscar la opción que corresponda a la pregunta que analiza sólo entre «a», «c» y «e». Finalmente, también es una prueba de lógica.

⁵² Zabalza remite a una experiencia de algunas escuelas inglesas, en las que los profesores amenazan con no recibir niños que no tengan padres ingleses, si se continúa afirmando que los resultados de la medición educativa dependen del trabajo docente

el aula, puesto que si el maestro considera que su función es llegar y explicar la clase (hablar), y la del alumno es atender a las explicaciones del maestro y estudiar, entonces parece que el examen sería un buen instrumento para detectar la información que retuvo. Pero si cambia nuestra concepción del aprendizaje, nos damos cuenta de que el alumno aprende en tanto se plantea problemas él mismo, elabora alternativas de solución y las puede discutir, es decir, si pone en juego una nueva información con respecto a su esquema referencial previo. Entonces la práctica de los exámenes resultaría inadecuada para estos planteamientos didácticos. Por otro lado, la preparación profesional implica el desarrollo de una madurez académica que va más allá del recuerdo de cierta información; más aún, en la práctica profesional lo que resulta relevante es la capacidad de resolver problemas, y no tanto la de recordar una información en un examen preestablecido. En esta perspectiva, el certificado de calidad profesional que expide el CENEVAL tiene un valor más simbólico que real para un desempeño profesional.

Es importante señalar que por la pretensión de proceder con *objetividad* en la estimación del aprendizaje, —pretensión que se mezcla con otros problemas, con buscar la *justicia en la evaluación*—, se han generado y difundido las pruebas «objetivas» o «estandarizadas»,⁵³ que consisten básicamente en presentar al estudiante una afirmación y pedirle que identifique o «adivine» la respuesta correcta de entre una serie de soluciones posibles. Los partidarios de este tipo de instrumento argumentan su «objetividad» en el manejo estadístico de la información, valor que obtiene a partir del empleo de un conjunto de técnicas derivadas de la estadística descriptiva, tales como la media aritmética, la desviación estándar, etc., estas técnicas no pueden dar cuenta de la parte sustantiva de la información que manejan: el valor del contenido, la importancia de la resolución de problemas, la capacidad de respuestas creativas, etcétera.

Entre las principales críticas que hemos planteado hacia el empleo de las pruebas tipificadas o estandarizadas, podríamos mencionar que este tipo de prueba difícilmente permite captar información que no sea de carácter memorístico. No negamos que existen ejemplos de reactivos que muestran lo contrario, pero precisamente éstos permiten confirmar las dificultades existentes para elaborar las preguntas y las limitaciones de estos instrumentos, los cuales, a su vez, impiden al estudiante plantearse problemas. Ya desde la década de los setenta, Vanheecke, bajo el título «Brasil: la computadora atonta la enseñanza», plantea cómo en Brasil han proliferado una serie de *cursinhos* para entrenar a los estudiantes en la identificación de las «mil y una mañas que permiten detectar la respuesta correcta en los tests de elección múltiple»⁵⁴

⁵³ Thorndike divide en cuatro partes la evolución de las pruebas objetivas o tipificadas: (a) 1900-1915, fase precursora de búsqueda y desarrollo inicial de métodos, basados en las escalas de bidet; (b) 1915-1930, auge del desarrollo de los tests, tests estandarizados para medir destrezas, baterías de tests de aprovechamiento, uso de ellos en la selección de personal para la *Army Alpha* y la *Woodworth Personal Data Sheet*, en este aspecto, a guerra también fué motor del desarrollo de las «ciencias del hombre»; (c) 1930-1945, periodo de evaluación, reconsideración, ampliación y delimitación, y (d) 1945-1960, periodo de baterías de tests y de los programas de aplicación de los mismos, en parte como consecuencia del éxito de su empleo en la Segunda Guerra Mundial. R. Thorndike y E Hagen 1970, pp. 13-16. Seguramente habría que añadir un último periodo, en el de internacionalización de las pruebas estandarizadas a partir de los exámenes nacionales de calidad profesional, propuestos en el marco de las políticas neoliberales en Occidente. Lo que cuevas denominó la derechización de Occidente a partir de 1975. A cuevas *et al.* 1986.

⁵⁴ C. Vanheecke 1975.

Tercera: la llamada evaluación con referencia a una norma (el desempeño de un grupo) y la evaluación con referencia a un dominio (desempeño frente al conocimiento) son, técnicamente, formulaciones sobre diferentes sistemas para asignar calificaciones

Un debate de los años setenta sobre la evaluación estableció la evaluación con referencia a una norma y con referencia a un dominio.⁵⁵ Himmel expresa que esta controversia se encuentra en los textos sobre «medición educacional⁵⁶ de los años setenta»; por su parte, Livas analiza lo que llama evaluación tradicional (con referencia a una norma) y lo que entiende por su transformación radical (evaluación con referencia a un dominio);⁵⁷ finalmente esta última es la que caracteriza la perspectiva de los exámenes nacionales de medición de la calidad. Sin embargo, antes de aceptar estas connotaciones, vamos a intentar caracterizar brevemente estas dos concepciones de la evaluación.

La *evaluación con referencia a una norma* presupone que las «aptitudes» son determinantes de un aprendizaje; las considera un pronóstico del rendimiento escolar, sin tomar en cuenta otros elementos que afectan el proceso del aprendizaje. Esta concepción considera que el aprendizaje se estima mediante la relación que se establece en un grupo entre el alumno que obtuvo la más alta puntuación y el que obtuvo la más baja. Estas dos puntuaciones se consideran los puntos de referencia para ubicar a los demás miembros de un grupo escolar, lo que se lleva a cabo mediante la comparación de cada puntuación obtenida con una muestra del grupo (que puede ser del 27% situados en el rango alto y otro tanto del rango bajo). Así, desde este punto de vista, lo importante es premiar a los «mejores» sin averiguar los elementos del proceso de aprendizaje que han repercutido en tal situación. Más aún, en el fondo poco importante si un sujeto ha aprendido, pues es más importante saber en qué lugar quedó con respecto al rendimiento del grupo. Mager caricaturizó esta situación cuando presentó el siguiente ejemplo: el piloto de un avión en vuelo se dirige a los pasajeros y les informa que es el egresado con las más altas calificaciones de la escuela de aviación; sólo lo habían reprobado en «aterrizaje».⁵⁸

Esta concepción con referencia a una norma se basa en la aplicación de la campaña de Gauss, la que se ha empleado para aventurar una «distribución normal de cualidades humanas, tal como se distribuyen en grupos numerosos tomados al azar»,⁵⁹ se trata de la aplicación de la teoría darvinista, que sostiene la sobrevivencia de los más aptos en un mundo de diferentes especies. Así, la evaluación es el mecanismo que permite que sobrevivan los alumnos que muestran mayores aptitudes. La inclusión de esta concepción en la evaluación privilegia fundamentalmente el manejo estadístico de datos, con la pretensión de lograr la buscada objetividad, la cual se limita sólo al manejo de los resultados, sin considerar que la verificación de un aprendizaje es algo muy diferente del manejo que se hace con tal dato. Ésa es la causa por la que los textos sobre evaluación dedican buena parte de su espacio al problema de una estadística simplificada, y quizá en ocasiones deformada. Vale la pena recordar que, en las tesis sobre evaluación que hemos

⁵⁵ El libro coordinado por Benjamín Bloom tuvo un papel clave en este debate. B. Bloom *et al.* 1975.

⁵⁶ E. Himmel 1979, p. 171.

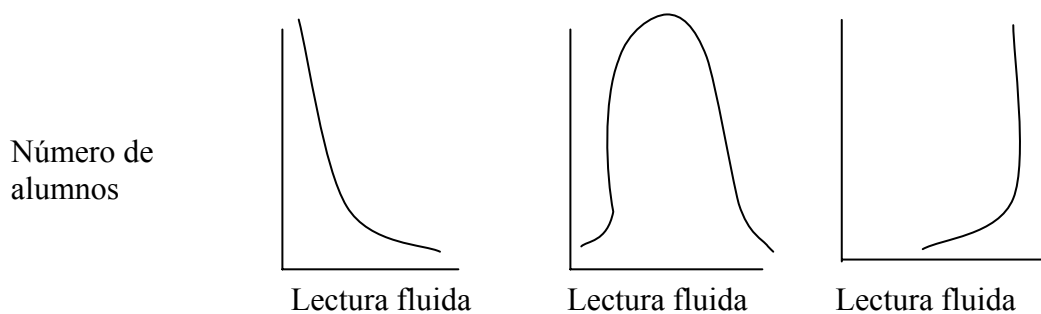
⁵⁷ L. Livas, 1978, p. 19

⁵⁸ R. Mager 1975, p. 104.

⁵⁹ G. Landsheere 1972, p. 96

establecido, mencionamos, de acuerdo con Benedito,⁶⁰ que la estadística utilizada en las ciencias del hombre no ha sido tomada directamente de las matemáticas, sino que ha sufrido mediaciones previas. Habría que considerar que la evaluación del aprendizaje no es un problema estadístico, y que se pueden establecer ciertos aspectos técnicos respecto a la manera como se emplea *la referencia a una norma*. Recordemos también que cuando más pequeña es una muestra, más se aleja de un universo para inferir una distribución «normal» de aptitudes.

En esta perspectiva, insistimos, subyace un biologismo social que pretende explicar las aptitudes únicamente con base en las leyes de la herencia y de la genética, presumiendo que el individuo posee o no ciertas aptitudes, a manera de «don» otorgado por la naturaleza. Es necesario mencionar que Landsheere dedica un capítulo al «Mito de la campaña de Gauss»,⁶¹ en el cual explica otros movimientos de esta curva normal de acuerdo con la distribución de frecuencia. Este autor establece un ejemplo tomado de la lectura, en el que una distribución de frecuencias en un grupo de alumnos de primer grado de primaria daría una forma de «y», pero si esa medición se efectuara en un grupo de alumnos de sexto grado, se encontraría que la mayoría de ellos puede leer, lo que permitiría identificar una curva con forma de «j». Esto lo podríamos representar gráficamente, como lo hace dicho autor, de la siguiente manera:



Capacidad de lectura

Fuente: G. Landsheere 1982.

Es necesario destacar que, dentro de las limitaciones de esta concepción, se puede incluir el planteamiento de otorgar las calificaciones a partir del desempeño grupal, no con relación a una «norma estadística», ni con la concepción de una «norma estadística», sino fundamentalmente reconociendo que el proceso del aprendizaje tiene particularidades que se expresan de forma individual y grupal. El valor y el significado de estas particularidades

⁶⁰ g. Benedito 1975

⁶¹ G. Landsheere 1972, pp. 190-202

en algún momento tienen que ser contruidos y valorados por los participantes de una experiencia de aprendizaje y no por elementos ajenos.

Por otro lado, la evaluación con referencia a un dominio o criterio es una manifestación diversa del interés de homologar los aprendizajes de los estudiantes y el valor de los certificados escolares. Cuando en los años cincuenta Bloom y sus colaboradores intentaban resolver por medio de una taxonomía la diversidad de comportamientos que se muestran en la acreditación, crearon paulatinamente el espacio para plantear la evaluación de dominio o «maestrías de aprendizaje». Esto significa que: (a) un estudiante puede aprender cualquier cosa, siempre y cuando se le den las condiciones y el tiempo para ello, y (b) la acreditación debía depender de que el estudiante dominase los contenidos propuestos en un plan y programa de estudios. En su perspectiva inicial, la evaluación con referencia a un dominio se tradujo en que cada estudiante manifestase el dominio o ejecución de los objetivos de comportamiento establecidos en un programa. De ahí que los programadores y los teóricos sugiriesen la elaboración de múltiples objetivos en un programa: generales o terminales, de unidad y, en ocasiones, hasta específicos. Estos objetivos representan únicamente fragmentos de la materia por aprender. Pager estableció los elementos que permiten «decodificar un objetivo»,⁶² con el fin de determinar el conjunto de acciones que el alumno tiene que realizar para acreditar un curso.

Esta perspectiva se vio considerablemente afectada por esa situación, pues el aprendizaje de un dominio también sucumbió en una trampa que lo vinculó con el conductismo, ya que se sigue concibiendo el aprendizaje como un producto que se puede «muestrear» por la cantidad de «ítems» contestados correctamente. Así, se reduce el aprendizaje a una dimensión lineal y acumulativa de información; por ello la información fragmentaria, que en forma de ejecuciones se desprende del objetivo, las más de las veces sólo puede ser memorizada. Con esta perspectiva es discutible —como lo sostiene Livas— que haya una transformación radical del problema

En este contexto, y para solucionar el problema aludido, se ha originado la idea de incluir los sistemas mecanizados, como la computadora y las pruebas «objetivas», para obtener del modo más eficiente la muestra de la información que se requiere en relación con los objetivos, o en su caso contenidos, a partir de un muestreo de información que no garantiza un aprendizaje. Los exámenes departamentales y los «nacionales de medición de la calidad» que han proliferado a partir de esta concepción, ponen énfasis sólo en el aspecto técnico de la evaluación, dejando de lado un concepto alternativo de aprendizaje.⁶³ Otro supuesto de esta concepción se vincula a la «maestría dominio en el aprendizaje», cuestión a la que Bloom dedicó sus esfuerzos en los años setenta; este punto de vista parte del supuesto de que «la aptitud es la cantidad de tiempo requerida por el alumno para alcanzar el dominio de una tarea inherente a su aprendizaje»⁶⁴. Con esta visión, es necesario que cada alumno dedique el tiempo que requiere individualmente para aprender. Debemos expresar que estos planteamientos guardan una similitud con los que fundamenta el sistema de

⁶² R. Pager 1975, p. 27

⁶³ A este respecto, Villarroel señala: «la acelerada tecnificación de la evaluación ha originado que se conciba a ésta como una estructura solamente psicométrica, divorciada casi totalmente del aspecto didáctico. Tal concepción nos parece errada y creemos que ha traído como consecuencia que los mayores esfuerzos y recursos para mejorar la evaluación se hayan dirigido hacia el problema del mejoramiento de las técnicas y no hacia el mejoramiento del proceso enseñanza-aprendizaje, como un todo, dentro del cual esta inmersa la evaluación». Véase C. Villarroel 1974.

⁶⁴ B. Bloom 1975, p. 77

Instrucción Personalizada de Séller, si bien frente a la pérdida del sentido social sobre la educación que la cosmovisión neoliberal ha efectuado en todo el mundo, quizá una mirada como la de «maestrías en el aprendizaje» resulta muy progresiva, pues ella no cierra la puerta de la educación a nadie, sólo pide que se le conceda el tiempo necesario para desarrollar su aprendizaje; mientras que la mirada neoliberal, al poner la eficiencia y la calidad sobre cualquier valor social, da pie a que se considere que sólo deben recibir educación quienes tengan cualidades —y posibilidades económicas, según el Banco Mundial—⁶⁵ para acceder a ella.

La evaluación con referencia a un dominio implica, a la vez, una escala de calificación dual, promovido o no promovido, mientras que en la referencia a una norma, lo importante es establecer una escala lo más diferenciada posible. Quizá valga la pena desarrollar otro elemento en esta concepción, el que se refiere a la acreditación, definida por el logro de todos los objetivos o el logro del desempeño complejo, de acuerdo con un análisis de tareas. No debe confundirse esta perspectiva con la llamada evaluación integral, que en ocasiones deriva en normas institucionales y prácticas docentes, en las que para la acreditación se busca tomar en cuenta una multiplicidad de elementos, tales como limpieza, orden, disciplina, actitudes de colaboración, disposición para el trabajo en equipo. Estos elementos, necesariamente formativos, son parte de la evaluación de un curso, pero no deben consignarse en la acreditación del mismo.⁶⁶ En este sentido, no cabe duda que el mérito de la evaluación con referencia a un dominio es efectuar la acreditación del alumno sólo a partir de los criterios de acreditación establecidos en el programa de un curso escolar.

Sin embargo, un tema central de nuestra exposición es que la llamada evaluación con referencia a una norma y la que se enuncia como evaluación con referencia a un dominio, constituyen formas diferentes de acreditación escolar, por medio de las cuales se otorga, con distintos procedimientos, la calificación obtenida en un curso. No son formas de evaluación, puesto que no permiten indagar acerca de los factores que afectaron el proceso de aprendizaje y las características de éste; sólo presentan diferentes opciones sobre la manera de calificar, y se limitan a buscar la posición del individuo con respecto a los demás integrantes del grupo, o bien plantear formas de derivar la calificación a partir del logro de un conjunto de comportamientos complejos exigidos como desempeño final de una actividad escolar. Estos planteamientos, en realidad, se preocupan por la «objetividad» y la «justicia» respecto de la manera de asignar calificaciones.

Sin embargo, este problema tiene que ser discutido a la luz de una teoría del conocimiento, puesto que la objetividad no se logra negando el factor humano⁶⁷ en el acto de conocer. En todo caso, habría que reconocer la presencia de lo «humano» como intrínseca al conocimiento del hombre. Entre otras disciplinas, los aportes teóricos del psicoanálisis brindan elementos para abordar este problema desde una óptica que permita incorporar elementos que apunten a una teoría de la subjetividad. Por ejemplo, las nociones de transferencia y contratransferencia aportan elementos útiles para otro tipo de

⁶⁵ Es evidente la implantación de una mirada economicista sobre la educación que se ha impuesto desde el Banco Mundial, y que ha sido refrendada

⁶⁶ También, respecto a la evaluación integral, Villarroel señala: «según la concepción de evaluación integral, un alumno puede ser promovido cuando no pudo adquirir la bibliografía porque es de escasos recursos, cuando tienen problemas familiares, estuvo enfermo, etc. Según nuestra opinión, no debe ser promovido; estos problemas del alumno no pueden compensarse con notas [...] los efectos de la evaluación integral han sido catastróficos, se pueden resumir en facilismo estudiantil». Véase C. Villarroel 1974.

⁶⁷ J. Bleger 1977, p. 114

comprensión y sobre el debate de la objetividad en la asignación de notas, ya que docentes y estudiantes se influyen mutuamente, y en su relación se mueven elementos inconscientes que vuelven a aparecer en el momento de determinar la acreditación.

La pretendida «objetividad» en la asignación de notas es un mito, puesto que, finalmente, el acto mismo de la elección de una calificación (aun cuando se encuentre protegido por elementos estadísticos o aritméticos, por ejemplo, la regla de tres) se encuentra marcado por la presencia de un *arbitrario*⁶⁸ que el docente puede manejar aun en la propuesta estadística. Este *arbitrario* justifica tomar alguna decisión que en los hechos modifica las notas de los alumnos. Así se puede favorecer o castigar a un grupo, de acuerdo con su desempeño previo. Sobre este elemento haremos algunas reflexiones más adelante.

Frente a la idea que busca objetividad en la asignación de la calificación, y precisamente por la presencia de este *arbitrario*, existen cinco formas básicas —y diferentes— de asignar una calificación, las que inicialmente podemos resumir de la siguiente manera: (a) dos se relacionan con la perspectiva de la *evaluación con referencia a una norma*, y (b) tres están vinculadas a la *evaluación con referencia a un dominio*.

Asignación de calificaciones con referencia a una norma

Hemos expresado que este tipo de asignación da pauta a dos formas de procedimientos que pueden emplear los docentes, una vinculada a la aplicación de elementos provenientes de la estadística descriptiva, y la otra cercana al empleo de una perspectiva aritmética, que en general termina asignando porcentajes a diferentes actividades de acreditación o empleando una regla de tres.

Manejo estadístico de datos. Esta opción se ha desarrollado inicialmente a partir de la generalización de las llamadas pruebas «objetivas»; cuando el docente las elabora y se trabaja con medios rústicos, se recomienda diseñar una batería que incluya, por lo menos, tres tipos de reactivos (correlación, opción múltiple, falso o verdadero, etc.), esta batería debe contener como mínimo 80 reactivos o «ítems». Sin embargo, conforme se ha ido introduciendo la computadora en la elaboración de este tipo de pruebas, la batería ha sido reemplazada por la creación de «bancos de reactivos» de opción múltiple, a los que se les aplica un procedimiento de «validación estadística», que en muchas ocasiones comienza sólo con un «juicio de expertos». De esta manera, se tiende a establecer y generalizar pruebas «tipificadas» o «estandarizadas», esto es, pruebas que responden a determinados valores estadísticos, tales como índice de dificultad o poder de discriminación. El valor estadístico de estos instrumentos no da cuenta de la certeza de la información que recuperan, ni del valor conceptual de la misma.

⁶⁸ Utilizamos la expresión *arbitrario* para expresar que, con independencia de la justificación racional sobre el procedimiento que cada institución o docente emplee para asignar la calificación, el hecho de que se haga una elección en la que se opte por un procedimiento y se excluyan otros —los cuales benefician o sancionan a los alumnos en sus calificaciones— hace que la calificación obtenida dependa también de los procedimientos elegidos. Así, con diferentes docentes un alumno puede obtener diferente calificación, pero un alumno con el mismo docente en diferentes cursos— el caso de los alumnos repetidores— también por ello puede variar su calificación. *Arbitrario* es una categoría conceptual que no ha sido suficientemente aceptada por la comunidad académica, desde que la trabajamos en 1982 en la revista *perfiles educativos*, o en 1984 en la primera edición de esa obra. En la decisión sobre el procedimiento para calificar a cada grupo escolar existen elementos que no se pueden explicar sólo por referentes teóricos, por ejemplo, los mecanismos de transferencia y contratransferencia. Es importante tener claro que *arbitrario* no significa capricho

La calificación estadística de los resultados que obtienen los alumnos en este tipo de pruebas se realiza a partir de procedimientos en los que se listan, en grado descendente, los puntajes obtenidos por los estudiantes que resolvieron el examen. Así, cuando una prueba tiene 80 reactivos, se parte de la mayor puntuación obtenida (supongamos 64) y las demás puntuaciones se van acomodando en grado descendente hasta llegar a la menor (supongamos 15). A partir de este ordenamiento, se empiezan a obtener la media aritmética, la desviación estándar, la moda, etc., para organizar distintos cortes de distribución que permitan formar o acercarse a una distribución normal de frecuencias. En esta distribución se asignan calificaciones de acuerdo con determinados cortes establecidos. Hasta aquí este procedimiento aparenta estar lejos de cualquier elección personal; sin embargo, como veremos más adelante, existen una serie de opciones técnicas que permiten favorecer o sancionar la calificación final que obtengan los alumnos, de acuerdo con lo que cada docente o responsable de esta calificación pretenda: beneficiar o sancionar.

Estas opciones estadísticas permiten modificar un poco el rango que puede abarcar cada intervalo de puntuaciones, y constituyen lo que hemos denominado *arbitrario*. Por ejemplo, el docente, o quien realiza el tratamiento estadístico de la información, puede eliminar la puntuación más alta y/o la más baja, en busca de hacer más compactos los intervalos de puntajes; con ello se busca una curva normal, ni puntiaguda, ni desvanecida. Un ejemplo evidente de este *arbitrario* se encuentra en los planteamientos que formula Ebbel,⁶⁹ quien sugiere que el docente elija entre una serie de adjetivos el que mejor le permita calificar el desempeño de su grupo. Estos calificativos se refieren a siete niveles de habilidad: excelente, superior, bueno, regular, promedio pobre o deficiente. Una vez que el docente elige el nivel de habilidad de su grupo. Obtiene en una tabla un valor numérico que le permite efectuar operaciones sobre la mediana de los puntajes organizados, y de esta manera asignar calificaciones. La elección de este calificativo es conceptualmente un *arbitrario*.

Una variante es la asignación de calificaciones utilizando procedimientos estadísticos, consiste en tomar en cuenta varias muestras de aprendizaje: pruebas, reportes de laboratorio, ejercicios de clase, ensayos, etc.; en estos casos también se puede asignar un valor estadístico. Por ejemplo, cuando en un curso se pide a los alumnos que elaboren tres trabajos monográficos, presenten dos exámenes parciales e integren diez reportes de laboratorio, a estos ejercicios se les puede asignar un valor, por ejemplo 30 puntos a los trabajos monográficos, 60 a las pruebas y 40 a los reportes de laboratorio. El máximo puntaje posible sería de 130 puntos; entonces se procede a organizar los distintos puntajes del más alto al más bajo, supongamos que de 98 a 32, y con estos datos se obtienen los diversos valores estadísticos (media, mediana, desviación estándar, moda) y se procede como se ha descrito para la asignación de la calificación.

Otro problema del empleo de la estadística descriptiva para la asignación de calificaciones, es que su uso no garantiza que exista una preocupación central por conocer el proceso de aprendizaje que tiene cada estudiante. En los hechos, este elemento pasa a un segundo término. Así se entiende la caricaturización que hace Pager cuando expresa: «Oh, muy bien, con este reactivo lograras una preciosa distribución de las puntuaciones. No sabrás si aprendió, pero la estadística será hermosa»⁷⁰ el empleo de la estadística descriptiva sólo vuelve más compleja la etapa de asignación de notas; a veces el docente —

⁶⁹ Para este tema véase f. García 1979, p. 166

⁷⁰ R. Pager 1975, p. 109

o quien realiza el manejo estadístico de datos— considera que mediante esta actividad efectivamente está haciendo una evaluación, aunque se pierde de vista el sentido de la misma; así subsiste la creencia de que la evaluación no es más que un conjunto de procedimientos estadísticos, y que su empleo confiere científicidad.

Manejo intuitivo de datos. Ésta es otra forma en que se puede emplear la concepción de la evaluación con referencia a una norma; en este caso no se procede a la conversión estadística, tampoco se intenta estudiar las causas que favorecieron o dificultaron el aprendizaje de los estudiantes. El supuesto que subyace en esta perspectiva es lograr que los alumnos den algunas muestras de aprendizaje, que no necesariamente tienen que relacionarse con los aprendizajes centrales de un curso; lo más importante, piensan algunos, es que el estudiante muestre disciplina en el trabajo entregando los reportes, tareas y exámenes que se le soliciten a lo largo del mismo. A veces en esta perspectiva subyace una visión de la llamada *evaluación permanente*, que defiende la idea de recolectar constantemente, a lo largo de un curso, un conjunto de datos. Pongamos como ejemplo el caso anterior: tres trabajos, dos exámenes y dos reportes de laboratorio; el docente distribuye la calificación entre este conjunto de datos, por ejemplo, 50% de la calificación por los dos exámenes, 30% por trabajos y 20% por los reportes de laboratorio.

En realidad no han ningún criterio que permita discutir si esta distribución de porcentajes es adecuada, dado que cada docente determina el valor de estas pruebas de la manera que considera pertinente; aquí vuelve a aparecer el *arbitrario* al que hicimos referencia en nuestra exposición. Esta forma de asignar calificaciones utiliza números absolutos, a los que les da un tratamiento aritmético a partir de la regla de tres; los docentes pueden fijar determinados criterios de aprobación, por ejemplo, aprobar los dos exámenes, tener un 80% de reportes de laboratorio, tener una calificación mínima en cada una de estas pruebas. Estos criterios se emplean para determinar si un alumno aprueba un curso; por ello en ocasiones se alejan de la posibilidad de comprender el proceso del aprendizaje del alumno, y más bien parecen ser la determinación de una serie de exigencias cuyo cumplimiento da la posibilidad de obtener una calificación.

Asignación de calificaciones con referencia a un dominio

Identificamos tres perspectivas para realizar esta tarea:

(i) *Integración de evidencias sobre todos los objetivos de comportamiento de un curso.* Antes que nada, deseo que quede claro que no recomendamos esta propuesta, puesto que lo que aquí se ha intentado explícitamente es hacer a un lado toda organización de programas de estudios centrada en la presentación de objetivos de comportamiento. Sin embargo, es necesario mencionar esta perspectiva, ya que es la que se desprende directamente de la denominada *referencia a un dominio*. Ahora bien, aplicar este enfoque resulta, en la práctica, sumamente difícil, pues se basa en la idea de calificar dominios que se derivan, en última instancia, de un «análisis de tareas», y cada tarea da como resultado un fragmento de información.

En este momento es importante aclarar que la fragmentación de comportamientos complejos para desarrollar habilidades particulares y posteriormente promover su integración, puede ser un instrumento de trabajo —incluso está previsto en la parte analítica de los métodos didácticos desde Comenio hasta nuestros días—; por ello, esta

fragmentación, apoyada en un análisis de tareas, puede ser muy importante en programas de desarrollo de habilidades psicomotoras. Sin embargo, el problema central de la propuesta comportamental es que no necesariamente todos los procesos cognoscitivos se desarrollan en un esquema de fragmentación, por el contrario, como sostienen la perspectiva gúestáltica y las propuestas cognoscitivas, los procesos de adquisición y construcción de la información requieren percepciones integradas y organizadoras avanzadas que permitan visualizar la estructura de una información y no sólo sus fragmentos.

Por ello, la propuesta de objetivos de comportamiento fragmenta tanto el contenido como el proceso de aprendizaje. En sentido estricto, la propuesta de evaluación con referencia a un dominio estableció que la acreditación sólo es posible cuando se dan pruebas de todos y cada uno de los comportamientos que establecen los programas a través de sus objetivos. Como un matiz de esta situación, en los años setenta se llegó a clasificar de manera arbitraria los objetivos de un programa en esenciales, necesarios y convenientes. A partir de esta jerarquización, se instrumentaron algunas pruebas para detectar manifestaciones de los comportamientos definidos como esenciales, otra para objetivos necesarios y otras para convenientes. Esta propuesta nunca llegó a aplicarse, pues su ejecución era muy complicada para docentes que atienden a muchos grupos y con programas llenos de especificaciones de comportamientos. Por ejemplo, si un grupo escolar tipo es de 40 alumnos y el programa de ese curso establece 60 comportamientos, el docente tiene que controlar en ese grupo 2 400 manifestaciones de aprendizaje. Este aspecto se complica si se considera que de esos 60 comportamientos, 20 fueron clasificados como esenciales, 30 como necesarios y 10 como convenientes. En esta propuesta, los alumnos que no dominan los objetivos esenciales no pueden acreditar un curso,⁷¹ esta posición nos remite de nuevo al *arbitrario* en la asignación de una calificación. Esto reclama que el docente elabore un informe pormenorizado de cada uno de sus alumnos sólo para especificar los comportamientos esenciales, para los cuales no existe manifestación de aprendizaje. La propuesta implicaría elaborar un programa de trabajo —el que finalmente tendría que tener un carácter individual adecuado para cada estudiante—, así como un programa de evaluación específico. Por demás esta afirmar que las condiciones de trabajo docente impidieron su aplicación.

Es necesario reconocer que la aplicación de pruebas y el manejo de datos mediante sistemas automatizados, como los que se hacen por computadora, pueden resolver este problema técnico, aunque no podemos perder de vista que estas pruebas, en general, trabajan sólo con preguntas que requieren un proceso de memorización, en detrimento de otras capacidades intelectuales como el juicio crítico, la creatividad, la capacidad de establecer relaciones, etc. en realidad, para resolver una prueba «objetiva», el estudiante necesita desarrollar otro tipo de lógica, como seleccionar entre posibles respuestas la que pueda ser más congruente o menos incoherente, y no plantearse forzosamente preguntas sobre un problema, ni sobre sus posibles soluciones.

Así, un problema fundamental de esta propuesta es la fragmentación y atomización del objeto de estudio; el muestreo de información de los programas de un curso de ninguna manera nos permite inferir que están en condiciones de integrar contenidos que han trabajado de manera fragmentaria y desigual.

⁷¹ «Se debe establecer previamente la puntuación aceptable, usualmente un 80% o 90%.» en este caso, tampoco se puede discutir cuál es el criterio para definir 80 o 90, lo cual es un arbitrario. I. Livas 1978, p. 54.

(ii) División de curso en etapas o secciones. Esta forma de asignar calificaciones se encuentra estrechamente relacionada con una perspectiva acumulativa de la evaluación. Consiste en dividir el programa de un curso en tres o cuatro partes; éstas pueden ser incluso las unidades formales del mismo. En general esta división obedece a un criterio temporal, de modo que se establece que en un semestre escolar⁷² tienen que existir cuatro calificaciones mensuales. En las instituciones donde se ha implantado el examen departamental,⁷³ su aplicación responde a un lapso en el que se supone que se ha cubierto, y por lo tanto se examinará, cierto tramo de programa.⁷⁴ En este caso se procede a obtener una calificación en cada etapa o unidad del curso, y la calificación final será el promedio de las anteriores. En un principio se puede pensar que a cada tramo o unidad del programa le corresponde un examen. Sin embargo, cada docente podría fijar el conjunto de pruebas que el alumno tiene que llevar a cabo para obtener la calificación en ese tramo, por ejemplo: realización de determinadas tareas o ejercicios, reportes de visitas, investigación, o prácticas y exámenes escolares. Según la perspectiva de la evaluación con referencia a un dominio, se puede establecer un mínimo aprobatorio en cada uno de estos bloques.⁷⁵ Esta situación, que es completamente racional e incluso recomendable en un curso escolar, no aleja el proceso de calificación del *arbitrario* que hemos enunciado.

(iii) *Asignación de calificaciones en relación con productos de aprendizaje*. Se apoya en la propuesta de elaboración de programas que hemos presentado en los capítulos anteriores.

Cuando establecimos la conveniencia de fijar, a partir de la precisión de una serie de *criterios de acreditación*, los requisitos para acreditar un curso, sostuvimos que éstos se presentan en dos niveles: una serie de requisitos formales que los estudiantes deben cumplir, y una serie de pruebas en el manejo de contenidos y habilidades que sean *significativas y pertinentes* para el proceso de aprendizaje, que impliquen algo más que la sola *memorización* de información y, sobre todo, que estén concebidas desde una perspectiva *integradora o estructurante* del contenido. En la definición de estas pruebas se pueden establecer las etapas de presentación de las mismas y su gradación en el desarrollo de un curso.

La experiencia nos ha enseñado que no es conveniente que no es conveniente que tales pruebas se exijan al final del curso, aunque puede ser muy práctico que los alumnos entreguen entonces el conjunto de actividades (trabajos, tareas, ejercicios, avances) realizadas durante el semestre. En la práctica, sobre todo cuando se pretende que los estudiantes efectúen procesos complejos, es muy conveniente establecer un mecanismo que

⁷² El *semestre escolar* no tiene la duración del semestre civil. Formalmente abarca seis meses, pero al quitar los periodos de vacaciones, de exámenes y de inscripciones habitualmente se reduce a cuatro meses de trabajo centrado en procesos de aprendizaje.

⁷³ El *examen departamental* es una prueba cerrada, tipificada; su estandarización se realiza a través de un análisis de reactivos que se hace generalmente por computadora. Se integran bancos de reactivos y el responsable del examen ordena su elaboración de acuerdo con los temas que deben haber sido cubiertos en un periodo escolar. El informe resultante es muy complejo: da información por tema, alumno, grupo e, indirectamente, da una «medida» de eficiencia docente.

⁷⁴ El artículo 6º del acuerdo señala: «La evaluación final será el resultado de la integración de las calificaciones parciales en cada materia o asignatura del plan de estudios, conforme a las disposiciones que establezcan las direcciones generales de la Secretaría de Educación Pública (SEP), que administren servicios educativos.» *Acuerdo No. 17*

⁷⁵ Véase Irene Livas 1978, p. 57. en lo que la autora denomina opción A, punto 3, establece la necesidad de fijar un mínimo por cada evidencia

permite ir revisando los avances del mismo a lo largo del curso escolar. Aun cuando consideramos que esta forma de realizar la acreditación tiene notorias ventajas, la definición de los criterios de acreditación y la interpretación de la información obtenida en cada prueba de aprendizaje no evitan la existencia de ese *arbitrario* que se presenta en la asignación de una calificación.

Para terminar este punto relativo a la determinación de una calificación, es necesario aclarar que el docente primero tiene que tomar una decisión sobre la acreditación, en la que determine el modelo de asignación de notas en el cual desea apoyarse; así puede identificar la relación entre pruebas de aprendizaje y contenidos (o en su caso habilidades) que se proponen en un programa del curso. Indudablemente, este segundo paso está influido por el *arbitrario*.

Es necesario considerar que en toda asignación de calificaciones interviene de una manera más o menos explícita un fenómeno de *transferencia y contratransferencia*, que hace que un docente experimente diversas sensaciones (aprobación, extrañeza, desconcierto) frente a los resultados obtenidos por los alumnos. Estos fenómenos no se resuelven evitando ver el nombre del estudiante cuando uno califica sus trabajos, en todo caso es importante que el docente genere una actitud de «autoanálisis» que le permita tomar conciencia de los mismos en estas situaciones.

En cuanto a la discusión sobre una calificación binaria (aprobado/no aprobado) o una de múltiples variaciones (4, 5, 6, 7, 8, 9 y 10), conviene tener presente que si bien las variaciones pueden dar una mayor precisión a la calidad del aprendizaje de los estudiantes, a la vez generan una serie de aspectos negativos en los mismos *estudiantes*, finalmente constituyen una etiqueta para ellos; en el *sistema de aprendizaje*, se estudia para tener una calificación no para aprender; en el *sistema social*, sólo se reciben quienes tengan determinada calificación de promedio. Para Pager,⁷⁶ la existencia de diversas calificaciones significa que el sistema educativo no está planificado para obtener éxito, sino para lograr fracasos. El pensamiento neoconservador promueve la diferenciación escolar —y la diferenciación social— en la actualidad a través de la generación de múltiples exámenes de «medición de la calidad»⁷⁷

La calificación no puede considerarse exclusivamente un problema técnico, ya que en lo fundamental es un problema social, tanto por las repercusiones individuales y sociales que tienen la certificación de conocimientos, como por los efectos distorsionantes que impone a los *padres de familia*, quienes ven en la calificación una gratificación al trabajo de sus hijos, los estudiantes, la mayoría de los cuales realiza su estudio sólo por la calificación, de tal suerte que puede decirse que se va a la escuela para obtener notas y no aprender.⁷⁸ Por su parte, la sociedad considera que las calificaciones son una segmentación de quienes tienen aptitudes para un trabajo y quienes carecen de ellas; mientras que *las*

⁷⁶ R. Pager 1975, p. 104.

⁷⁷ En el caso mexicano resulta muy interesante la información que filtra el CENEVAL, a los medios de comunicación, relativa a los resultados de los exámenes nacionales de ingreso a la educación superior, por ejemplo: «la mayor calificación fue de 89.44% y correspondió a Rubén Alberto Díaz Castaño de Chihuahua», mientras que en el Examen General de Calidad Profesional de Contaduría, «la mayor calificación fue de 77.73 aciertos de Patricia del Carmen Marrero de Yucatán». Véase I. Melgar 1995, pp. 1 y 11. Esto convierte la evaluación en un acto histórico, y se falta a la ética en el manejo de la información.

⁷⁸ Giner de los Ríos cita cómo los maestros en la Universidad de Oxford se quejaban a mediados del siglo pasado de que, con la aparición de las calificaciones, los alumnos habían perdido el placer de aprender y ellos el placer de enseñar. Véase Giner de los Ríos 1906.

instituciones y los responsables de la política educativa exhiben las buenas calificaciones como si fuesen equivalentes al esfuerzo educativo, en detrimento de todos los esfuerzos que se realizan para trabajar con quienes tienen mayor dificultad de aprendizaje. Hoy es más fácil, desde el punto de vista de la medición nacional, expulsar a un alumno que tiene un rendimiento deficiente que enfrentar el reto pedagógico de trabajar con él.

En todo caso, valdría la pena precisar que los textos que abordan el tema de la evaluación no lleguen a descubrir ni a estudiar el *arbitrario* que subyace tras el conjunto de opciones técnicas que se usan para determinar una calificación de los estudiantes. Un estudio de la evaluación desde la perspectiva del valor y la función social de la nota escolar, y de las repercusiones individuales en relación con las motivaciones del sujeto, así como sobre los mecanismos de transferencia y contratransferencia que afectan esta tarea, contribuiría a desarrollar una *comprensión*, en la tradición de la teoría del conocimiento, que permitiese desarrollar otra perspectiva respecto del proceso de la evaluación del aprendizaje y, por supuesto, de la acreditación escolar.

Cuarta: considerar la asignación de notas a partir de una teoría de grupo representa un reto para la educación.

El momento de la asignación de notas es uno de los más conflictivos, tanto para los docentes como para los alumnos, por cuanto implica que el maestro asuma el papel de *juex* y en ocasiones de policía, para decidir sobre la calificación de un alumno, sobre el valor académico y fundamentalmente social del estudiante, por las connotaciones que tiene la calificación. También es cierto que, por una determinación institucional, el docente no puede negarse, a realizar esta actividad. En este sentido, una alternativa para enfrentarse este problema es conceder alguna participación en la asignación de notas a los alumnos y al grupo escolar. Esto se hará siempre y cuando exista un proceso de trabajo durante el semestre que por su seriedad, compromiso académico y manejo de la información de los contenidos lo permita. Para ello, el docente debe apoyar su trabajo en una concepción de grupo. Cuando hablamos de grupo, debemos referirnos a distintas tradiciones, desde la que se ubica en la perspectiva de los grupos en relación con la eficiencia (Lewin), hasta la que aporta la teoría psicoanalítica de los grupos (Foulkes, Bion, Anzieu), y al desarrollo latinoamericano, que ha tenido cierta aplicación en nuestros medios académicos, desde la perspectiva operativa de grupo o en sus desarrollos institucionales.

Con fundamento en una concepción grupal se pueden generar diversas prácticas educativas en el aula, lo que va más allá de un empleo mecánico de las diversas técnicas grupales, o de la promoción y realización de trabajos en «equipo», que terminan siendo una expresión de la división técnica del trabajo (cada alumno hace aisladamente una parte de un trabajo más amplio); estas dos últimas situaciones han dañado la enseñanza, han viciado a los alumnos y afectan notoriamente las posibilidades del empleo de una perspectiva grupal en el trabajo en el aula. En este aspecto deseamos reiterar que, sólo a partir de una formación en grupo y de la adopción de una concepción grupal, puede el docente intentar desarrollar una propuesta en su salón de clase.

Algunos autores⁷⁹ hablaron de una *didáctica grupal*, aunque debemos reconocer que, con el surgimiento de los colectivos sociales, paulatinamente el trabajo en grupo se fue

⁷⁹ C. Hoyos 1980, p. 21.

incorporando a la práctica escolar, en especial con los desarrollos de las distintas propuestas de la escuela activa, en particular la de Freinet. A partir de una perspectiva grupal en la didáctica se pueden establecer formas metodológicas de enseñanza que cambien el papel de receptor del alumno, o el papel de quien realiza sólo las tareas de asignar el maestro, a un papel de compromiso, búsqueda de información, discusión y colaboración horizontal. Sólo cuando se han transformado las condiciones de aprendizaje de los estudiantes —cuando éstos muestran un compromiso con su proceso de aprender— puede avanzar la participación del grupo en el proceso de evaluación.

En este punto resulta relevante la experiencia de la escuela activa, en particular el papel que Freinet⁸⁰ asigna a los compañeros de una clase para elaborar felicitaciones o críticas a las conferencias, textos, reportes de bibliotecas que presentan los alumnos de estas escuelas ante sus compañeros. Esta experiencia muestra las posibilidades de una evaluación que no se limita a la asignación de notas. Algunas veces, cuando el proceso grupal ha sido excepcional por el papel que los estudiantes han tenido en un curso —es decir, cuando se ha generado un proceso de aprendizaje de su papel en la escuela, propiciando que desarrollen un proceso de autoobservación, que analicen lo que significa este proceso, reconociendo lo que se va logrando y la cantidad de problemas, contradicciones, lagunas y dudas que una nueva información plantea—, se pueden realizar algunas experiencias sobre la participación de los estudiantes en una primera asignación colectiva de calificaciones.

Dos aspectos habría que cuidar en esta actividad: (a) que se hayan entregado las diferentes pruebas de aprendizaje solicitadas en el curso (ejercicios, reportes, ensayos, etc.) y que su presentación sea pública, esto es, que los estudiantes hayan tenido acceso a este material; (b) es muy conveniente que profesor tenga una calificación de cada estudiante y que la muestre una vez que de forma individual o grupal se haya asignado una calificación al grupo, de suerte que se puedan ver los puntos de coincidencia y de discrepancia y analizar las causas de los mismos.⁸¹

Lo importante, desde esta perspectiva, es considerar que si se desea modificar los procesos de asignación de calificación, es indispensable modificar los procesos de enseñanza. Sólo transformando la enseñanza se puede mejorar la educación (lo que hoy llaman calidad de la educación) y sólo así se puede mejorar el sistema de evaluación. No es con exámenes más rígidos o más extensos (aplicados en todo el país) como la educación se mejora. También es importante considerar que la evaluación es un espacio de experimentación de alternativas; el conjunto de ponencias presentadas en el *VI Encuentro Nacional de Escuelas alternativas*⁸² muestra que es factible salirse de una tradición

⁸⁰ Véase C. Freinet 1973.

⁸¹ En mi trabajo profesional, en muy raras ocasiones he podido realizar este tipo de experiencia, la cual es muy significativa para los estudiantes y para el mismo docente. Se trata de experiencias serias, realizadas con alumnos que muestran cambios actitudinales significativos, en las que entre varios hemos leído sus trabajos. En estas experiencias hemos visto que no existe una gran discrepancia entre la calificación (y motivos de asignación de la misma) que se asigna grupalmente y aquella que en mi papel de docente he propuesto. Obviamente, esta calificación la llevo por escrito a esa sesión, pero sólo la muestro una vez que se ha asignado la calificación grupal. Véase A. Díaz Barriga 1994^a, pp. 140-145.

⁸² En general responde a un movimiento que adopta diversos principios del movimiento de las escuelas activas, incorporando nuevos elementos en su conformación. Se presentaron diversas ponencias sobre los sistemas de evaluación del aprendizaje en la Escuela Paidós, Bartolomé de Cossio, Centro educativo Pátzcuaro, Escuela Freinet de Cuernavaca, Instituto Escuela, etc. el evento se realizó en Pátzcuaro en febrero de 1995.

enciclopedista y examinadora, y es hacia donde caminan las propuestas pedagógicas de la modernización educativa de los años noventa.

Por nuestra parte, seguimos comprometidos con generar diversas formas de evaluación. La participación del grupo en la asignación de notas es el término de un proceso iniciado en las primeras sesiones de un curso, y mantenido con enorme esfuerzo tanto del docente como de los alumnos a lo largo de cada sesión de trabajo. Estas experiencias nos invitan a seguir desarrollando nuevas aproximaciones a este problema, pero debemos cuidar su difusión para que no degeneren en un facilismo pedagógico o en un acto irresponsable por parte de maestros y alumnos. Somos testigos de que esta propuesta también ha dado pauta a que, ante la carencia de una propuesta sólida y consistente de aprendizaje, se recurra a este expediente como respuesta fácil de un maestro que sólo desea quedar bien con sus estudiantes. Nos oponemos a esta situación.

No es fácil romper con una inercia que lleva a recurrir en la acreditación a la mentira y al ocultamiento de un proceso por parte de docentes y alumnos. Los procesos de asignación de notas deben abandonar el juego de «policías y ladrones», en el que el docente tiene que «vigilar» al que «copia» o mostrar que un trabajo es reproducción de otro. Pro superar este juego implica modificar profundamente los conceptos y las formas de enseñanza. La evaluación es un resultado, no es el punto de partida para la innovación. Por eso desconfiamos de los exámenes de medición de la calidad y de los sistemas nacionales de evaluación. En poco tiempo estos sistemas se encontrarán saturados de «folletos» que presenten reactivos sobre los cuales se examinará, e incluso de pequeños señalamientos para lograr mayor eficiencia al responder este tipo de instrumentos. Mientras tanto, los alumnos continúan desarrollando una manera de engañarse a sí mismos, según la cual lo importante es obtener una calificación y no reconocer el proceso de aprendizaje que tienen.

Es fácil desarrollar vicios en la propia percepción sobre los logros obtenidos, es más cómodo depositar la responsabilidad de todo en un docente, y mejor aún, en un sistema anónimo que aplica pruebas, entrega resultados e «histeriza» con ellos a la sociedad;⁸³ mientras, los estudiantes mantienen los mismos vicios, incrementan la dificultad para realizar autoevaluaciones, para mirarse de cara a sus propios resultados.

Reconocemos que son muy pocas las situaciones en las que un grupo escolar puede participar completamente en la asignación de sus notas, no así en la evaluación general del curso y en el desarrollo de una mirada crítica tanto de sus logros como de sus tareas pendientes. Sin embargo, en algunas de estas experiencias los alumnos han podido manifestarse críticamente frente al conocimiento, frente a su aprendizaje y frente a su calificación, y presentan esta manifestación ante su grupo. Lejos de lo que se puede pensar, en estas experiencias la generalidad de los alumnos opta por no asignarse notas altas, y nuestro problema, como coordinadores, ha consistido en ayudar al grupo a pensar en su aprendizaje y a evitar la complicidad que se puede generar por no reflexionar sobre las experiencias de aprendizaje que ha vivido.

Deseamos insistir que esta experiencia sólo se puede realizar —y en muy pocos grupos— cuando ha existido un proceso grupal previo en el que los participantes han roto con múltiples esquemas (vicios) preestablecidos. Un síntoma puede ser muy elocuente de ello: si los alumnos no se han comprometido apasionadamente con una bibliografía que vaya más allá de lo que incluye el programa y con el desarrollo extraordinario de los

⁸³ Véase Y. Stronach y B. Moréis 1994, pp. 5-19

productos establecidos en los criterios de acreditación, es imposible que realicen esta experiencia.

Sin embargo, en aquellos grupos en los que experiencias de este tipo no se puedan realizar, sigue siendo válida la distinción entre evaluación y acreditación, puesto que de todas maneras es conveniente y posible efectuar una evaluación sobre el proceso desarrollado desde el principio del curso al momento actual. Esta evaluación tiene como tarea que cada integrante y el grupo reflexionen sobre su proceso de aprendizaje, sobre las dificultades que tuvieron que afrontar y la manera como las fueron abordando, los resultados a los que llegaron y las carencias que perciben.

Indudablemente, efectuar la evaluación es un paso necesario para poder dejar en manos del grupo la acreditación. Dicho de otra manera, si en la evaluación grupal los coordinadores perciben que el nivel alcanzado por el proceso del grupo les permite responsabilizarse de su acreditación, ésta se puede llevar adelante. Es aquí, y no al principio del curso, cuando se puede decidir si el grupo está en posibilidades de asignarse una nota.

No hay que perder de vista que, por las condiciones sociales ya interiorizadas en cada sujeto de manera no consciente, aparecen en este momento un conjunto de resistencias, conflictos, evasiones y complicidades que si no se trabajan en el momento de su aparición, impiden la realización adecuada de esta experiencia, la que, sin embargo, es una posibilidad muy importante para sacar al docente de su papel tradicional de *juez*.

Para finalizar, diremos que hemos querido externar en estas líneas las dificultades que se presentan hoy en día en relación con el discurso actual de la evaluación, generado en la perspectiva pragmática y completamente asociado al ascenso del conservadurismo de Occidente.

Por ello hemos querido plantear algunos elementos teóricos que permitan generar un enfoque distinto para esta actividad, así como presentar algunas posibilidades prácticas diferentes. Ésta no es una formulación teórica que dé respuesta acabada a los problemas planteados, sino más bien un punto de partida para la construcción de una teoría de la evaluación, y una la reflexión sobre algunos aspectos que de ella se derivan para la docencia.

SANTOS, Guerra Miguel A. “Evaluación cualitativa de centros escolares” En: La evaluación: un proceso de diálogo, comprensión y mejora. Granada Ed. Albije pp. 109-138

CAPÍTULO IX

EVALUACIÓN CUALITATIVA DE CENTROS ESCOLARES

EN EL UMBRAL

La actividad de los Centros escolares no se suele someter a una evaluación sistemática y rigurosa (Simons, 1987; Santos Guerra, 1989). Las escuelas, sencillamente funcionan. Cada año, cada día, se repiten las mismas rutinas sin saber muy bien por qué, y para qué. La evaluación se realiza en ellas (no sobre ellas) para comprobar si los alumnos han *aprendido* lo que se les ha pretendido enseñar.

El Centro es una *caja negra* en la que entran los alumnos con un caudal de saberes muchas veces no se sabe muy bien cuál es), y de la que salen con otro bagaje. No se conoce lo que pasa dentro de ella, cómo se produce el aprendizaje y qué tipo de procesos se producen en la actividad interna.

Esta *caja negra* no se encuentra suspendida en el vacío sino que se sitúa en un contexto socio/económico/político amplio (en una sociedad) y en un contexto próximo (en una zona o barrio) en el que desempeña un papel. Dentro de ella se desarrollan unos mecanismos para producir fuerza de trabajo y unos procesos de aprendizaje de las relaciones sociales de producción.

“En enseñanza, hay aspectos que van directamente enfocados a la formación de la fuerza de trabajo, a aumentar su productividad (términos de valor) y otro que simplemente, satisfacen una demanda social de acceso a la cultura, en el sentido cotidiano que damos a este término”

(Fernández Enguita, 1985).

En el Centro escolar cumple una misión social con independencia de lo que cada individuo condiga en él. En realidad no es fácil dilucidar si la organización está al servicio de los individuos o si son estos los que, a través de ella, facilitan que se preceptúe un determinado orden social.

Reducir la evaluación de un Centro a la preocupación por los resultados académicos de los alumnos, encierra una simplificación considerable. Si los alumnos *brillantes*, si los estudiantes *sobresalientes* han aprendido también en la escuela a ser opresores, a ser insolidarios, debería preocuparse por haberles dado tantas armas...

Los Centros escolares desempeñan una tarea en la sociedad y no puede reducirse su evaluación al simple cómputo de aprobados y suspensos.

“El aprendizaje de la constitución –antes leyes fundamentales-, la religión o la historia, es decir, de esa constitución, esa religión, y esa historia, es mucho más una necesidad de la sociedad de cara al mantenimiento del orden –de ese orden- que una necesidad de los individuos, pero la sociedad se las arregla para convertirla en tal”. (Fernández Enguita, 1985).

El *currículum* oculto que se desarrolla en el Centro, la “política de etiquetado” (Apple, 1986), el aprendizaje de normas y patrones de conducta, nos hacen considerar la evaluación

institucional de una escuela como un proceso ambicioso que trata de responder a esta pregunta: ¿Qué está sucediendo aquí?

Cada Centro trata de dar explicaciones, casi siempre gratuitas, a lo que sucede dentro de sus muros. Y establece conexiones de causalidad interesadas. Así, los padres pueden atribuir el fracaso de sus hijos al mal funcionamiento del Centro, el Director puede valorar positivamente el trabajo porque muchos alumnos aprueban la selectividad, los profesores pueden achacar el fracaso de los alumnos a la poca aplicación y a la falta de capacidad de los estudiantes o de los profesores procedentes.

Los procesos atributivos, nacidos frecuentemente de la parcialidad y el interés, no permiten reconstruir la realidad, comprenderla adecuadamente, ni por supuesto, impulsar un cambio efectivo. Cada persona, atribuye las causas del éxito o del fracaso a aquello que no cuestiona su compromiso con la institución o su concepción de la misma.

En la *evaluación informal* de Centro se utilizan criterios excesivamente simplificados. ¿Cuántos alumnos aprueban?, viene a ser la pregunta central. El porcentaje de aprobados se da como resultado objetivo y se explica de forma también simplista, aunque interesadas. No se plantea en el análisis en qué condiciones se ha aprendido, qué tipo de alumnos han estudiado en el Centro, qué evaluación se ha realizado, qué alumnos han abandonado previamente, qué tipo de familias tienen sus hijos en el Centro, etc.

Cuando se realiza una *evaluación formal externa* se suelen aplicar pautas e instrumentos que difícilmente puede reflejar con rigor toda la complejidad de los procesos que se desarrollan en el Centro (y que, a su vez, tienen complicadas conexiones de input, output y throughput con el exterior).

Este tipo de evaluación, realizada a través de cuestionarios aplicados a varios profesores o directivos del Centro, no puede recoger su verdadera realidad ya que:

- ☉ Es imposible encerrar la extraordinaria complejidad de la vida de un Centro escolar en las encorsetadas respuestas de un cuestionario.
- ☉ La discordancia entre el pensamiento y acción es también preocupante. Lo que se hace es diferente a lo que se piensa sobre aquello que se hace. Un Director puede pensar que sus comportamientos son democráticos. Otra cosa muy distinta son las actitudes y los actos a través de los cuales se relaciona con el profesorado y los alumnos.
- ☉ Puede existir discordancia entre lo que se dice a unos y lo que se diría a otros. No es lo mismo contestar un cuestionario sobre el Centro para la Inspección que para una entidad que va a conceder una subvención.
- ☉ Existe, casi necesariamente, una discrepancia entre lo que dicen unos y lo que dicen otros. Hacer una evaluación del Centro a través de la aplicación de un escaso número de instrumentos valorativos a pocas y determinadas personas de la misma institución extraña un riesgo de arbitrariedad y parcialidad ilimitada.
- ☉ Existe también discordancia entre lo que se dice en unos medios y en otros. No se valora igual al Centro a través de una entrevista que por medio de un cuestionario. De ahí la debilidad metodológica que supone hacer una evaluación institucional únicamente a través de la contestación a las preguntas de un cuestionario, si éste es superficial y, sobre todo, sino tiene carácter anónimo.

- ☞ Existe discrepancia entre lo que se dice en un momento y en otro. Por eso no es fiable una evaluación que se realiza en un corte transversal muy corto de tiempo, necesariamente fugaz. En un momento determinado puede existir en el Centro un conflicto que está condicionando toda la apreciación que se hace sobre él.

La evaluación interna (auto evaluación institucional) suele hacerse en España a instancias de la Administración, que ordena que se presente anualmente una

MEMORIA DE CENTRO.

- ☞ En realidad no se trata de una auto valoración propiamente dicha ya que es una imposición burocrática. De ahí que la Memoria se haya convertido en un trámite sin repercusión alguna sobre la toma de decisiones. Se hace la Memoria académica de un curso, y al siguiente, se comienza a funcionar como si no se hubiese realizado.
- ☞ Solamente intervienen en la realización de Memoria los profesores, de modo que nada se sabe a través de ella sobre el pensamiento de los alumnos y de los padres sobre el Centro. Cuestión decisiva ya que en una evaluación democrática tienen que intervenir los protagonistas de la actividad.
- ☞ Sólo participan personas que están dentro de la institución. Existe el riesgo de que la subjetividad condiciones el rigor del análisis.
- ☞ La memoria es, frecuentemente, un simple catálogo de actividades. Se dice todo lo que se ha hecho, pero no se analiza cómo se ha realizado, por qué se ha hecho y que resultados ha tenido.
- ☞ Cada profesor realiza su Memoria del curso de forma autónoma e independiente. No suele afrontarse la tarea desde una perspectiva conjunta, interdisciplinario, multifacética y sistemática.
- ☞ La Memoria se realiza en tiempos “libres” del profesor, hurtados a otros menesteres o a su tiempo de ocio. No se reserva un tiempo al desarrollo de la evaluación institucional. Parece lógico destinar a la acción muchas horas. La evaluación, al no ser considerada un elemento decisivo de mejora institucional, es realizada de forma anecdótica, circunstancial y precipitada.
- ☞ La Memoria se realiza al finalizar el curso, de manera que se convierte en un modelo de evaluación sumativa, no realizado durante la acción y, por consiguiente, está más pendiente de resultados que de procesos.

Los profesores tienen conciencia de que la Memoria es un simple trámite burocrático. No puede ser leída ni analizada por las instancias superiores que la solicitan y la imponen. En definitiva, se trata de una exigencia de la Administración que desvela dos grandes problemas: la desconfianza de que el profesorado realice por propia iniciativa una tarea fundamental y la falsa vía de ordenancismo como forma de mejorar la actividad profesional. Si la evaluación institucional es importante, habrá que conseguir que la formación y el perfeccionamiento del profesorado sean tales que esta actividad surja de forma espontánea y sea realizada con tiempo y preparación suficientes. Habrá que hacer posible que las plantillas se formen por criterios pedagógicos y no exclusivamente por intereses personales o derechos adquiridos.

SIMPLIFICACIONES ABUSIVAS.

La tradicional forma de evaluar los Centros escolares ha estado presidida en nuestro país por fórmulas excesivamente simplificadores. Basta revisar la bibliografía al respecto (García Hoz y Medina Rubio, 1987; Moreno García, 1978; Martínez, 1979; Isaac, 1983;

Tort, 1985; etc.) para darse cuenta de que una curiosa particularidad: SE explicita la complejidad de la tarea y se propone facilitar extremadamente simplificada de llevarla a cabo.

Las pautas cuantificables que tratan de “medir” el funcionamiento del Centro y de “comparar a través de la medida” están muy lejos de poder facilitar la evaluación del Centro Escolar desde una dimensión ambiciosa.

Si se tiene en cuenta, además, que esas pautas se aplican a directivos, profesores y padres de alumnos, podremos hacernos una idea de su “peligrosidad”, ya que, aunque la muestra sea representativa, los protagonistas no pueden agotar los puntos de vista sobre el Centro. Cuando la demanda del proceso evaluador viene de instancias exteriores al Centro, es fácil que los “actores” intenten proyectar una imagen del Centro positiva. Al pretender comparar dos Centros a través del resultado cuantificador de un cuestionario, se comente una arbitrariedad ya que:

- ☉ Los contextos en que se desarrollan las actividades son distintos. Una escuela enclavada en una zona rural no es fácilmente equiparable a un gran Centro radicado en el corazón de una urbe. Una escuela suburbial no es comparable con un colegio al que acude la élite social. No hay Centros, pues, estrictamente comparables. ¿Sería razonable comparar la velocidad de dos coches de distinta cilindrada y atribuir la diferencia a la pericia del conductor? ¿Sería justo comparar la velocidad de dos vehículos similares puestos a circular por firmes tan diferentes como una autopista y un desierto?
- ☉ Los medios disponibles son diferentes. No es riguroso comparar, a través de instrumentos idénticos, los resultados de potenciales diferentes alimentados por medios muy distintos. Distintos en cantidad y en calidad. Volviendo al ejemplo del coche, diríamos que no tiene sentido comparar los kilómetros recorridos o la velocidad alcanzada por dos coches que tienen diferentes cantidad y calidad de combustible para la marcha.
- ☉ Los patrones de medida son utilizados de forma diferente. Cada uno contesta al cuestionario desde postulados diferentes y a ceses, hasta opuestos. Los conductores serían en este caso los encargados de manifestar a qué velocidad han recorrido y cuántos kilómetros ha sido en trayecto.

El tema se plantea en la entraña misma de la comparación. ¿Se puede evaluar sin disponer de un criterio comparativo? Stufflebeam y Shinkfield (1987) adoptan una respuesta ecléctica ante este interrogante.

“Creemos que la evaluación debe ser comparativa antes de que empiece el proceso o antes de suscribirse a un servicio, no comparativa durante el proceso (para que permita concentrarse en lo que resulte mejor para el desarrollo del servicio), y periódicamente comparativa, después del proceso y/o utilización continuada, para que el camino quede abierto ante posibles mayores avances”.

La pretendida científicidad de los números no hace más que enmascarar una realidad compleja. El tratamiento estadístico posterior funcionará con la inflexibilidad de los programas matemáticos, pero difícilmente podrá reflejar la realidad vida, compleja, multiforme, imposible de encerrar guarismos. Una realidad que sólo adquiere sentido en la interpretación contextualizada de sus códigos y en la atribución de valor y de mérito a sus

diversas facetas. De ahí que una misma forma de actuar (por ejemplo, en la disciplina existente en un Centro) pueda ser valorada de formas muy diferentes.

En cierta ocasión tuve que presentar una película en un Centro de Madrid, los alumnos interrumpieron su actividad ordinaria para asistir a la proyección. Entraron con tal orden y silencio en la sala, a un metro de distancia unos de otros, con las manos atrás... que no pude reprimir la sorpresa y la desaprobación. Para los profesores del Centro aquella forma de actuar era, sin embargo, la deseable.

Es difícil encerrar este patrón de comportamiento en un cuestionario. Y es difícil interpretarlo sin atender a códigos contextualizados que permitan atribuir sentido y valor a los hechos.

El intento de cuantificar el complejo entramado de relaciones, acciones y significados está amenazado de trampas e imprecisiones. Es probable que, en estos comentarios, incurramos en el mismo defecto simplificador que denunciarnos. El objetivo de estas líneas es llamar la atención sobre algunas de las limitaciones de los intentos de medición del Centro escolar, sobre todo cuando se aplican de forma exclusiva, sin ningún otro tipo de método o instrumento que facilite el contraste, la matización y la profundización en los datos obtenidos. Los cuestionarios pueden ser *útiles*, pero encierran el peligro, cuando se aplican con carácter *exclusivo*, de recoger de forma superficial y reduccionista una realidad extremadamente compleja.

Cuando nos referimos a los diseños experimentales como forma de explorar la realidad y de construir el conocimiento, no pretendemos hacer una descalificación de sus bases científicas y de sus estrategias metodológicas sino subrayar algunas de sus limitaciones, especialmente relevantes cuando se trata de analizar una realidad social tan compleja, dinámica y cargada de valores como la de un Centro escolar.

La polémica *métodos cualitativos versus métodos cuantitativos* resuelta hoy por los especialistas de ambos campos, con un intento de aproximación, comprensión y conjugación metodológica. (Cook y Reichard, 1986). Sin embargo, la naturaleza de algunos fenómenos sociales aconseja un enfoque cualitativo, ya que su complejidad no puede ser captada plenamente a través de diseños experimentales (Stake, 1983; Pérez Gómez, 1983b; Angulo Rasco, 1988).

Las mediciones

Las escalas de estimación, los cuestionarios, las pautas de evaluación que exigen valoración numérica, no dejan paso a la explicación causal de la realidad. Supongamos que la disciplina de para un profesor encuestado 4 sobre 10 para un profesor encuestado. No explica por qué atribuye ese número ni por qué es esa la realidad del Centro. “Cuatro” no “significa” nada. Para una persona, un orden riguroso, un silencio estricto, una sumisión absoluta, etc. Resultan indicadores de un planteamiento disciplinar equivocado. Para otra persona son esos patrones los que merecen la puntuación máxima. Por otra parte, una misma puntuación puede ser asignada por motivos justamente contrarios.

El fenómeno de la atribución numérica a realidades complejas supone un riesgo importante, no sólo de impresión, sino de tergiversación, agravado por el hecho de la apariencia de objetividad que encierra el número. Un número parece más objetivo, más riguroso que una descripción expresada en términos naturales.

“Nuestra búsqueda de la racionalidad científica y la evidencia factual nos ha llevado a una reverencia irracional por la autoridad de los números raramente cuestionada”. (Siedman, 1977).

Los cuestionarios son contestados, como es lógico, por los “actores_”, por los protagonistas de la acción educativa en el Centro. Este hecho añade un factor de complejidad al fenómeno de la medición. Un director a quién se pregunta que ponga una calificación numérica al nivel de información que tienen los alumnos en el Centro, estará fácilmente inclinado a responder con una calificación alta. En parte, también, porque, aunque reconozca con precisión la realidad, tenderá a presentar una imagen favorable del Centro que dirige.

La cuantificación obliga a encerrar en los números realidades cargadas de matices y de complejidad. Constreñir una organización tan complicada como es el Centro escolar, en la que se entremezclan factores ideológicos, políticos, sociales, económicos, psicológicos, pedagógicos, en un conjunto de números, encierra riesgos evidentes:

- a) En el proceso codificador, ya que se parcela una realidad única y se califica numéricamente un hecho cualitativo, desde la particular subjetividad del calificador.
- b) En el proceso decodificador, ya que se atribuye un significado, quizás desde otros códigos, a los números asignados por el calificador.

Olvidarse de este proceso de atribución y comenzar a barajar los números y a operar estadísticamente con ellos (Más directamente, a atribuir un valor comparativo a los guarismos) es un ejercicio descarado de arbitrariedad, con escasas apariencias de rigor para quien se detenga a considerar, desde el comienzo, el proceso evaluador.

La calidad de los datos es discutible por el proceso de atribución que supone el asignar un guarismo a una realidad en sí misma complicada y contemplada desde la subjetividad del participante.

Las comparaciones.

Las escalas que se utilizan no sólo suelen ser nominales (7 es diferente de 6, 4 es distinto de 5) sino que tiene el carácter de ordinales (7 es más que 6, 4 es menos que 5) e incluso se les atribuye una condición abusiva de escalas de razón (7 menos 6 es igual a 5 menos 4).

Esto permite fáciles comparaciones y clasificaciones. Un Centro que obtiene una media de 7’5 en la valoración de la limpieza es “mejor” que otro que solamente obtiene un 5’5. (No importa de qué medios de limpieza dispone, qué personal se dedica a la tarea, en qué lugar se encuentra ubicado el Centro, qué tareas realizan los usuarios, qué climatología existe en la zona, qué tipo de alumnos lo integran...) Ni siquiera se tiene en cuenta el criterio de exigencia de los evaluadores (profesores o alumnos del Centro), de tal manera que “los más limpios” –y, por consiguiente, los más exigentes—resultaría ser “los más sucios” en a evaluación, ya que otros evaluadores (profesores y alumnos de otro Centro), menos exigente, otorgarían puntuaciones más altas en limpieza del Centro a pesar de estar más sucio.

Estamos eligiendo un ejemplo del menor grado de complejidad posible, pero la cuestión se puede ir complicando a medida que los indicadores que se tengan en cuenta estén referidos a cuestiones más complejas, como el grado de libertad de los alumnos, el nivel de

participación, la existencia de indoctrinación, el respeto a los individuos, la naturaleza y niveles de los aprendizajes, etc.

Utilizar las pautas cuantificadas como criterio de comparación, resulta, pues, arbitrario. No es acertado considerar como Centro “mejor” (¿Mejor para quién?, ¿mejor para qué?, ¿mejor en que circunstancias?, ¿mejor en que aspectos?, ¿mejor durante cuanto tiempo?...) a aquel que consiga puntuaciones más altas en una evaluación cuantitativa.

Las interpretaciones

El lector/investigador debe descifrar lo que aparece cifrado en los informes. El acto de codificar encierra una operación formal consistente en compendiar, en reducir muchas cosas a una. Sólo comprendiendo las claves de esa operación, se puede realizar un correcto descifrado. Cuando no se dispone de las claves o de las reglas, hay que realizar la operación a través de conjeturas.

¿Cómo se pueden conocer unas claves que el sujeto no ha desvelado? Cada evaluador contesta desde su peculiar criterio a los cuestionarios, atribuye números sin precisar cuáles son las reglas críticas mediante las cuáles ejerce esa atribución. El lector, que tiene acceso directamente al número y no a la realidad, difícilmente podrá descifrar la entraña de aquel número. Es probable que la arbitrariedad se potencie, ya que está en el camino de ida y –

REALIDAD ————— NÚMERO ————— RECONSTRUCCIÓN

además –en el de vuelta: ¿Qué significa un 3 que ha atribuido un profesor al estado del mobiliario escolar de un Centro? ¿Ha puesto una calificación baja porque el mobiliario está viejo, porque es inadecuado, porque se usa mal, porque le disgusta su color, porque es poco manejable, porque está en mal estado, porque no tiene estética, etc.?

Añadamos la compleja gama de posibles reacciones del profesor en la dinámica del Centro: El Director ha impuesto el tipo de mobiliario, el tenía especial interés en otro tipo de material, la Administración ha servido los pupitres iguales a todos los Centros, etc.

Sin conocer al profesor, sin saber cuáles son las características del Centro mal se podrá realizar una interpretación ajustada.

Cuando se aplican las pautas numéricas, sin tener en cuenta el contexto, sin disponer de las claves de interpretación, la falta de validez del instrumento es inevitable.

Los paquetes estadísticos que se apliquen una vez obtenidos de *esta manera* los números tienen dudosa Validez, ya que parten de unos datos arbitrarios.

La realidad queda plasmada en un número que le interpreta, la condensa, la resume. A través de este número, otro lector reconstruye la realidad. A nadie se le oculta que la naturaleza de ese filtro de doble sentido ofrece una inevitable dificultad para la fiel interpretación de lo que ocurre en el Centro. Este filtro no sólo es insuficiente, sino que es fácilmente deformante.

Añádase a esta complejidad intrínseca el problema del sesgo intencionado. Al evaluador directo (profesor, director, alumno) le interesa que la “imagen” de esa realidad sea buena. Y

también le puede interesar que sea positiva al evaluador que recaba los datos. La identificación con un tipo de Centros, con su ideología, con su emplazamiento, con las personas que han accedido a realizar la evaluación, etc. Hace que la aplicación de estos cuestionarios cerrados, que filtran la realidad a través de números, resulten sumamente problemáticos.

La interpretación puede referirse también a las medidas estadísticas extraídas de diferentes medidas parciales. Varios profesores evalúan, por ejemplo, el funcionamiento de los órganos colegiados. Sus puntuaciones serán, lógicamente, diferentes. También lo serán las de los alumnos. ¿Qué sentido tiene hallar la media aritmética de las puntuaciones? ¿Qué interpretaciones puede darse a una puntuación extraída de criterios tan diversos y quizás opuestos?

Lo mismo decimos de la interpretación de las diferencias cuando se evalúan comparativamente varios Centros. ¿Qué se dice *realmente* (no sólo estadísticamente) al afirmar que existen diferencias “significativas” entre dos Centros? La extensión de la significación estadística a otras vertientes (sociológica, pedagógica, didáctica, etc.) es claramente abusiva.

Las generalizaciones.

Cuando se aplican diseños de carácter experimental, sean causales, correlacionales o “ex post facto”, se suelen realizar generalizaciones estadísticas. La naturaleza estocástica de las mediciones permite aventurar unos resultados similares en un nivel de probabilidad determinado.

Esta generalización de los resultados a otras realidades parece justificar una buena parte de la investigación. La validez externa del experimento radica precisamente en esta potencialidad que permite aplicar a otros casos resultados hallados en la muestra con que se trabaja.

Uno de los problemas de la generalización estriba en la endeblez de la extracción de la base numérica que ha servido para aplicar el modelo estadístico. Ya hemos visto con qué dificultad puede asignarse un número a una realidad compleja, analizada subjetivamente por el evaluador.

Otro problema es la peculiar idiosincrasia de los Centros. No es fácil (o al menos, no es riguroso) parcelar la realidad de los Centros. Cuando se hace así a través de los números y se opera luego estadísticamente con ellos, se descontextualiza el dato, y se le priva del significado.

No tiene sentido afirmar que la cultura de los bantúes es *significativamente diferente* a la cultura de los mayas. Son realidades distintas, absolutamente “incomparables”. Decir que la edad a la que mueren los bantúes es *significativamente distinta* a la edad a la que mueren los individuos mayas tiene muy poca relevancia. O, mejor dicho, tiene una relevancia que sólo puede ser extendida cuando se estudia y se tiene en cuenta todo el complejo entramado de ambas culturas.

(Nada decimos aquí de la dificultad de elegir muestras representativas de realidades tan complejas. Cada Centro puede ser entendido solamente desde sus propias coordenadas.

OTRA DIMENSIÓN: LA EVALUACIÓN ETNOGRÁFICA.

El propósito principal de la evaluación, tal como aquí la entendemos, es conocer como y porqué funcionan de una manera determinada los Centros, para comprender y explicar su naturaleza (Formulación de teorías) y para mejorar su práctica, enriqueciendo la toma de decisiones.

Para mejorar la práctica es preciso conocerla en profundidad. Reconstruir la realidad desde una perspectiva crítica no sólo supone su contemplación aséptica y su descripción rigurosa. Es necesario interpretarla, saber dónde están las raíces de los comportamientos, las causas de las actuaciones, los efectos del desarrollo curricular (tanto explícito como oculto).

En el Campus de la Universidad de Norwich los alumnos de Lawrence Stenhouse plantaron un árbol como homenaje póstumo. En la placa situada al pie del árbol escribieron en esta sentencia clave de la obra del autor inglés: “Son los profesores los que, al fin, podrán cambiar el mundo de la escuela, comprendiéndolo”. Es la comprensión profunda del fenómeno educativo que se produce en la escuela lo que provocará la transformación y el cambio en profundidad.

No importa mucho llegar a conclusiones genéricas del tipo “Este Centro funciona muy bien”. Y menos al realizar afirmaciones como “Este Centro es mejor que este otro”. Nos interesa analizar la realidad del Centro, saber qué fenómenos se desarrollan en él, por qué se producen, que repercusiones tiene, etc. Para ello necesitamos:

1. Tener en cuenta la importancia del contexto.
 - 1.1. el contexto diacrónico, porque el Centro no puede entenderse sin su historia. El Centro es lo que es hoy, funciona como funciona, porque ha tendió en cuenta los antecedentes inmediatos, un conflicto no puede comprenderse sin acudir a las causas que lo han provocado.

En un Centro escolar privado que conozco el titular expulsó en el mes de julio a 30 de los 40 profesores que componían la plantilla. ¿Podría comprender un evaluador externo el ambiente del Centro, al comenzar el curso siguiente, sin tener en cuenta lo sucedido, cuando ese hecho estaba condicionando absolutamente toda la realidad?

- 1.2. el contexto sincrónico porque los códigos de un Centro (lingüístico, comportamentales, ideológicos, etc.) son diferentes. El contexto estará determinado por elementos diversos, unos de carácter interno (configuración del Claustro, Proyecto Educativo, extracción social de los alumnos, etc.) y otros de carácter externo (emplazamiento, clase social de las familias, naturaleza del medio, etc.).

“la conducta humana, dice Brofenbrenner, está influida significativamente por los ambientes en los ambientes en los que ocurre” (Brofenbrenner, 1976).

2. Enfatizar el valor de las relaciones e intercambios de carácter psicosocial.

En los intercambios se cimenta toda la acción educativa. El Centro escolar está atravesando por un entramado de relaciones interpersonales de diverso tipo: de

autoridad, de sumisión, de enseñanza/ aprendizaje, de información, de transacción económica, etc.

Estas relaciones se establecen en el marco del aula, pero también fuera de ella. No se puede circunscribir toda la comunicación a la esfera del aula, ya que en el Centro se establecen muchas relaciones, algunas de las cuales condicionan las que se producen en el aula. Y dentro del aula, no todas las transacciones o intercambios se desarrollan sobre contenidos de carácter formal ni, desde luego, sobre contenidos de aprendizaje reglado.

Todo el entramado de relaciones conforma el clima institucional en el cual se desarrolla la acción.

*Las imaginaciones que la gente tiene de los otros, de ellos y de cada aspecto de sus mundos son los **hechos sólidos** de la sociedad. Percibiendo su mundo, el hombre es subjetivo. Selectivamente percibe, interpreta y atribuye significados en su mundo y, entonces, actúa en consecuencia. Los **hechos sólidos** de la sociedad, son esas percepciones, definiciones, interpretaciones y significados” (Dodge, y Bogdan, 1974). (El subrayado es mío).*

3. interpretar representativamente los hechos.

No basta la descripción minuciosa de los hechos, no basta un simple análisis. El factor clave dentro de la comunidad social es la interpretación representativa de los fenómenos.

La interpretación representativa trasciende y supera la simple enumeración y descripción de los fenómenos. Y supone un análisis que maneja símbolos y valores.

“la investigación interpretativa, entiende la conciencia de una situación más allá del mero conocimiento proposicional, al ámbito de los sentimientos, al ámbito de los sentimientos, de las simpatías silenciosas, a los deseos inconscientes y a los usos cotidianos no examinados, profundizando nuestra comprensión de los ambientes organizativos y sociales”, apunta Vax (1971). Y añade:

“un investigador de campo que explora un pueblo extraño percibe enseguida que la gente está diciendo y haciendo cosas que ellos comprende, pero que él no. un sujeto puede efectuar un gesto por el que todos sus compañeros se serían. Ellos comparten la comprensión de lo que el gesto significa, pero no así el investigador. Cuando él también lo comparte comienza realmente a comprender”.

Se hace necesaria, pues, la interpretación representativa de los hechos, porque éstos constituyen un lenguaje cargado de significaciones.

4. Subrayar el valor de los procesos, no sólo (aunque también) en de los resultados.

Importa tener en cuenta cómo se hacen las cosas, cómo se desarrolla la actividad, cómo se relacionan las personas, cómo se organiza el complejo mundo de los profesores, cómo se relacionan las personas, cómo se organiza el complejo mundo de los profesores, cómo se desarrolla el proceso de enseñanza/aprendizaje. Cómo se toman las pequeñas y grandes decisiones, etc.

No nos importa tanto cuántas cosas se hacen, cuanto alumnos aprueban, cuántas decisiones se toman, cuántos éxitos se consiguen, etc. No es que no importen los resultados conseguidos, sino que la forma mejor (quizá la única) valorarlos debidamente es saber cómo se producen.

5. Utilizar la opinión de los protagonistas de la acción.

Aunque plateamos la necesidad de realizar una evaluación externa del Centro escolar, será preciso contar con la valoración de los profesores, de los alumnos y de los padres que forman la comunidad escolar.

Son ellos los que realmente conocen los hechos por dentro y los que dominan las claves para su interpretación, la cultura del Centro puede ser descifrada por los miembros de la misma.

Si los protagonistas mantienen una actitud diferente o de rechaza hacia el proceso evaluador, difícilmente se podrá obtener una información fidedigna u, desde luego, de poco servirán los resultados de la evaluación para mejorar la profesionalidad de los profesores y para influir positivamente en la posterior toma de decisiones. Ahora bien, esa negativa o indiferencia son una parte sustancial de la institución.

6. Mantener una visión estructural del Centro

El análisis parcelado de una realidad compleja presenta vetas muy profundas de distorsión. No se puede entender, por ejemplo, lo que sucede en un aula sin tener en cuenta las características de la organización general del Centro, no se puede entender la actuación de un Director sin conocer la actitud de un conflicto entre dos grupos de profesores sin saber cuáles son las causas que lo han desencadenado, etc.

El centro es un todo y no se puede evaluar por parcelas sin riesgo de que a explicación quede distorsionada al descontextualizarla. No sólo porque unas cosas dan sentido y completad a otras sino porque un fenómeno aislado tiene, incluso, distinta significación en un contexto y otro. Es probable que la discrepancia de un profesor ante la dirección del Centro tenga un diferente significado en un Centro privado y en uno público, por ejemplo.

Cuando, al salir de un túnel, leemos en el cartel indicador APAGUE LAS LUCES, debemos tener en cuenta el contexto en el que se sitúa la lectura; si ese día, haremos bien en apagar las luces del vehículo; si es de noche, deberemos hacer caso omiso de la advertencia.

Una vez conocido el contexto, podemos realizar un ENFOQUE PROGRESIVO, centrándonos en un aspecto determinado en el que nos interesa profundizar. Es el efecto "zoom" o de traveling óptico" aplicado al análisis de la realidad. Imaginaremos que, al realizar la evaluación del Centro, nos encontramos con un conflicto de autoridad. Si deseamos profundizar en su análisis, polarizamos la atención y el esfuerzo en el estudio de sus causas, de sus dimensiones, de sus consecuencias, de su duración, de su profundidad. Pero sólo podremos comprender el fenómeno si lo enmarcamos en sus coordenadas etiológicas y de desarrollo.

Puede suceder que el zoom tenga sentido inverso (que se comience a estudiar un fenómeno determinado del Centro, porque el interés sea exclusivo o porque la gravedad del problema lo aconseje). También en ese caso será preciso ir ampliando el marco de referencia para comprender la realidad que despertó el proceso de evaluación.

7. Dar prioridad a la vertiente cualitativa de la evaluación.

La naturaleza de los fenómenos que ocurren en el Centro escolar es de carácter social. De ahí que utilicemos enfoques, método, instrumentos y análisis de carácter cualitativo. Las objeciones que se suelen formular a este planteamiento (“no hay ciencia de lo particular”, “no hay rigor en lo que es subjetivo”, “no existen criterios de contraste”, “no hay posibilidad de generalizar”, “no existe la necesaria asepsia ideológica”, etc.) han de ser resueltas a la luz del paradigma de investigación utilizando. No parece razonable estudiar las características de un proceso de investigación desde las condiciones y exigencias de otro. En cualquier caso, habría que exigir al proceso aquellas condiciones que le hagan riguroso, coherente y adecuado a la realidad que se explora.

“si en evaluador en busca de la verdad ignora estos hechos (subjetivos) a través de la imposición de sus definiciones en lo que él cree que sea el ambiente de investigación, está ignorando datos valiosos” (Dodge y Bogdan, 1974).

Es frecuente que se intenten justificar la bondad del paradigma cualitativo en la evaluación / investigación desde las premisas y las exigencias del paradigma cuantitativo. Lo cual encierra un error, cuando menos, de perspectiva. Y, probablemente, de valoración, ya que parece existir un paradigma fundamental y modélico en el que han de mirarse los demás. El criterio de bondad y de rigor de los demás sería la medida en que se pareciesen al modelo.

Cook y Riechardt (1986) han realizado una aportación interesante al respecto, planteado la compatibilidad de los métodos y sugiriendo la conveniencia de su utilización adecuada y, en algunos casos, conjunta.

Es preciso elaborar mecanismos de información son suficiente capacidad de rigor para conocer qué es lo que está sucediendo en el Centro, cómo se está desarrollando el currículum y qué es lo que debe transformarse para la mejora.

Este sentido se pronuncia Gimeno Sacritan (1998):

“Un sistema que no dispone de mecanismos de información sobre lo que produce queda cerrado a la comunidad inmediata y a la sociedad entera, sin la posibilidad de que ésta en su conjunto, previamente informada, pueda participar en su discusión y mejora. La política educativa, la evaluación de la validez de los *curricula* vigentes, la respuesta de los centros a su comunidad quedan sin contrate posible; los mismos profesores se justifican con acomodarse a la regulación abundante a la que es sometida su práctica. El *currículum* que no se evalúa, o se hace a través de la evaluación de los profesores solamente, es difícil que entre en una dinámica de perfeccionamiento constante. Sin información sobre el funcionamiento cualitativo del sistema escolar y curricular, los programas de innovación o mejoras pueden quedarse en la expresión de un puro voluntarismo o en sometimiento a iniciativas que podrían no responder a necesidades reales del sistema escolar, de los alumnos y de los profesores.

8. Crear hábitos y actitudes que favorezcan la autorreflexión la autorreflexión sistemática y rigurosa.

La evaluación que proponemos es siempre, en buena medida, auto evaluación. Cuando es externa porque necesita de la participación inexcusable de los protagonistas que son consultados por los evaluadores. Y, cuando es interna, porque son los mismos sujetos los que se preguntan y responden por la reconstrucción de la realidad del Centro.

En cualquier caso, para que la evaluación se convierta en un proceso de mejora de la realidad ha de acabar siendo asumida como un elemento capaz de incidir en la toma de decisiones en el Centro. La evaluación impuesta o sufrida resultará estéril. Y el conocimiento de la realidad que se puede conseguir a través de ella ser posible y desvirtuado.

Los informes que no se solicitan o que molestan serán arrinconados, manipulados o denotados por aquellos que han sufrido la imposición de ser evaluados.

Despertar y cultivar actitudes de autorreflexión es conseguir poner en marcha procesos de evaluación continua y sistemática. Hará falta, solamente, añadir el carácter técnico y riguroso que exige una evaluación explícita y que se materializa en informes.

MÉTODOS DE EXPLORACIÓN.

No existe un camino único para llegar a reconstruir la realidad, para conocerla, para explorarla e interpretarla.

Cook y Reichardt (1986 son partidarios, como hemos dicho anteriormente, de utilizar aquellos métodos que sean más pertinentes con el proyecto, dando validez a los experimentales (cuantitativos) y a los etnográficos (cualitativos).

La realidad educativa es tan compleja que no se puede abarcar su comprensión con un solo instrumento de exploración. Vamos a plantear algunos criterios para la selección y utilización de los métodos.

Adaptabilidad.

El conocimiento de los sujetos y de la situación aconsejará la elección de unos métodos u otros. Y el momento y la forma mejores para su aplicación. Se exige, pues, una permanente actitud de apertura y una sensibilidad extremadamente atenta a las indicaciones y exigencias del medio en el que se investiga. El diseño ha de ser, pues, flexible.

En lugar donde el uso del video sea excesivamente distorsionador (pensemos en un Centro en el que los alumnos no están familiarizados con las cámaras) no será conveniente comenzar haciendo grabaciones comprometidas.

De lo que decimos se deduce que, para elegir lo métodos, no sólo importa la naturaleza de los fenómenos que se exploran (¿Qué es mejor, un artillo o un serrucho?) sino el momento y el lugar en el que se utilizan.

Variabilidad.

Una de las formas de conseguir datos fiables es utilizar instrumentos de distinta índole que permitan el contraste y la compensación de las limitaciones de los otros.

Así la observación de un aula deberá ser contrastada con entrevistas a profesores y alumnos y con la aplicación de cuestionarios. Utilizar un solo método encierra peligros de parcialidad y de sesgo intrínsecos al mismo e impide la posibilidad del contraste, del análisis cruzado...

Además, los datos obtenidos a través de un método pueden ser utilizados para su profundización /constate/ matización a través de otro; por ejemplo, los datos recogidos en la observación pueden servir de base para el desarrollo de la entrevista. El evaluador, a través de diversas preguntas, podrá comprobar el sentido de lo que ha observado o de lo que ha leído en un cuestionario o en un documento oficial del Centro.

Gradualidad.

No se deben utilizar las técnicas con la misma intensidad en todo momento y circunstancia. En parte porque el evaluador no dispone de los datos necesarios para profundizar en los diferentes ámbitos y temas de la realidad, en parte porque es probable que los protagonistas no estén dispuestos a abrir plenamente los reservorios de las esencias del significado.

Una exploración inicial puede abrir paso a otras posteriores de mayor profundidad. Así, una entrevista hecha al comienzo puede ser más breve y superficial que otra realizada más avanzada el estudio, cuando existen más datos sobre los que trabajar y condiciones más favorables para la profundización en los temas.

Pertinencia.

El conocimiento de la realidad y de las técnicas aconsejará la utilización de alguna en momentos determinados según la naturaleza de los datos que se pretende obtener, de los fenómenos que se quieren explorar o de las personas que van a ser exploradas.

Es, probablemente, más difícil conseguir información valiosa sobre lo que piensa el profesor a través de una grabación de video que mediante una entrevista o de la técnica de pensamiento en voz alta.

Dominio.

Hay evaluadores que tienen un especial dominio o habilidad en el manejo de algunas técnicas. Está claro que esas serán perfectamente utilizadas en razón de la gran cantidad y riqueza de los datos que puede obtener a través de ellas.

Un evaluador especializado en la observación obtendrá una gran variedad de datos significativos que no le permitirá conseguir la entrevista si se siente incómodo en el uso de esa técnica.

Los diferentes evaluadores pueden trabajar preferentemente con aquellas técnicas con las que se sientan más familiarizados, procurando que estén debidamente compensadas.

La importancia de la figura del evaluador en este tipo de trabajo es decisiva. Digamos que todo descansa sobre su preparación científica, sobre su claridad en la utilización de los métodos y sobre su habilidad y honestidad para relacionarse con los integrantes de la comunidad escolar.

La experiencia, la edad, el sexo, la compostura, el tacto del evaluador son aspectos que hay que tener en cuenta para la elección y la utilización de los diferentes métodos de exploración de la realidad.

Eric Hoyle (1986b), en un sugerente trabajo, dice que “las técnicas de evaluación no pueden ser especificadas para todas las escuelas”. La elección dependerá, añade, de lo que ha de ser evaluado, de las características de la escuela y de los evaluadores y, también, de las bases sobre las que se han realizado la negociación entre la escuela y los evaluadores.

La observación persistente, la entrevista en profundidad, el análisis de documentos, el diario del evaluador, etc. son técnicas que se utilizan de forma cruzada para obtener una información de contraste que depure el significado de lo que se pretende evaluar. Evaluar con validez supone y exige la aplicación rigurosa de métodos de exploración de la realidad (House, 1980).

EL INFORME Y LA RETROALIMENTACIÓN.

El informe final es de exclusiva competencia y responsabilidad del evaluador (y de su equipo). Un mal informe invalida un buen proceso de evaluación.

Resulta esencial garantizar la independencia del evaluador en el momento decisivo de redactar el informe. Hay que proteger al evaluador, hay que fortalecer su papel con el fin de garantizar la libertad:

- ♣ Ante el patrocinador (Administración, Titular del Centro, etc.). ya que tiene fuerza para presionar, condicionar, censurar e, incluso, eliminar el informe.
- ♣ Ante los miembros de la comunidad escolar que, a veces, desean recibir valoraciones halagadoras, positivas y gratificantes, eludiendo o descalificando aquello que resulte desagradable.
- ♣ Ante la sociedad, que necesita valoraciones rigurosas que orienten sus elecciones, que posibiliten la mejora de la calidad de las instrucciones que ofrecen un servicio tan importante como la educación.

El evaluador ha de ser claro y preciso en el informe. Para ello ha de buscar un lenguaje asequible al destinatario. Un lenguaje arcano, cargado de expresiones científicas, técnicas o artificiosas, hace inútil el trabajo. Los interesados no tienen acceso al descifrado de mensajes crípticos. La pretendida científicidad que se quiere hacer surgir de los tecnicismos lingüísticos esconde la verdadera riqueza del discurso y entra en la contradicción, en este caso, con la naturaleza etnográfica del estudio.

La redacción no ha de ser aburrida, difícil y árida, ya que el destinatario es quien lo lee, no quien lo escribe. Los códigos no son siempre los mismos para el emisor que para el receptor. El evaluador debe tener en cuenta que el informe va a ser leído por personas que son necesariamente expertas en materia educativa.

El rigor de las conclusiones ha de estar basado en la riqueza y precisión de la indagación. Por eso, las conclusiones han de aparecer avaladas por la descripción metodológica que ha precedido a su elaboración. Las conclusiones no aparecen por generación espontánea en el marco del informe.

Schwartz y Jacobs (1984) hablan del método de las *confesiones verdaderas*. El autor expresa, a través de él, cómo y porqué se condujo en la evaluación. Sin embargo, dicen, esta pretendida fidelidad “no resuelve el problema que nuestros métodos de observación y análisis de la realidad no pueden funcionar en el mundo social”. Esto es cierto, pero no lo es menos que la valoración de las conclusiones dependerá del rigor de la exploración.

Si el evaluador ha querido confirmar sus propias hipótesis, de forma más o menos sutil, más o menos consciente, doblegando los hechos al compás de sus intereses, de sus teorías apriorísticas, resultará difícil hacer llegar al destinatario un razonamiento riguroso que muestre la fuerza del discurso y de sus corolarios y conclusiones.

Puede haber diferentes tipos de informes:

- a) Según el destinatario.

Para adaptarse al destinatario hay que realizar informes de extensión, contenido y naturaleza peculiar a sus características. No es que haya que modificar el contenido, pero sí la forma y la extensión.

- b) Según la finalidad.

La divulgación exige un informe sintetizado de lo que resulta más relevante y significativo. La investigación exige un informe minucioso, extenso, presidido por el rigor.

Cuando hablamos de conclusiones nos referimos a las afirmaciones que se derivan del trabajo, al estilo de las conclusiones estadísticas que confirman o rechazan las hipótesis previas. En el desarrollo del trabajo podrían encontrarse explicaciones de la realidad que, de alguna manera, pueden considerarse conclusiones parciales.

En la evaluación, etnográfica son los mismos lectores que formulan sus propias conclusiones, deduciendo interpretaciones de las descripciones, de los comentarios, de las entrevistas y de los registros de acción.

Es preciso apoyar/ subrayar las interpretaciones en las evidencias de los datos, aunque, como es lógico, el evaluador tendrá que realizar una *lectura entre líneas*, tendrá que recoger el sentido debajo de la capa superficial de los hechos. Las apariencias en un Centro no siempre permiten desvelar a primera vista lo que realmente sucede en el mismo.

La redacción del informe requiere que el evaluador desarrolle destrezas de descripción. Su capacidad de observación, su intuición, su sagacidad y perspicacia en la interpretación de los hechos, pueden esterilizarse ante la incapacidad narrativa.

Además de la pericia, hace falta mucho tiempo.

Elliott (1987) plantea una secuencia temporal que puede servirnos de referencia:

- 1 día para la observación.
- 4 días para redactar el informe.
- 2 días para la negociación.
- 1 día para la revisión del informe.
- 1 día para la nueva negociación.

La entrega del informe (en las condiciones que se precisaron en la negociación) no ha de demorarse mucho, en parte por que la dinámica del Centro varía con el tiempo y en parte porque se pierde curiosidad e interés por los resultados de la evaluación.

“la utilización (del informe) podría darse en dos niveles. En primer lugar, en la comprobación de la exactitud de los datos. ¿Ha producido u informe riguroso de este acontecimiento? ¿Están bien representadas ciertas impresiones? ¿Se han tomado en cuenta todos los aspectos pertinentes? En segundo lugar, sobre cualquier interpretación o explicación, el informante puede tener algún comentario que hacer” (Woods, 1987).

En el informe no deberían incluirse datos que no tengan, al menos, dos fuentes contrastadas, de lo contrario, el evaluador se arriesga a que falte objetividad en la apreciación.

Para potenciar la objetividad ante el lector, es conveniente presentar, al final del trabajo, un apéndice con las pruebas testimoniales de donde se han extraído los datos (registro de la observación y de la entrevista, documentos de diverso tipo, cintas de video, fotografías, etc.).

El informe de la evaluación de un Centro no debe contener críticas y descalificaciones personales, no debe emitir juicios de valor sobre la actuación de los individuos, sino descripción de situaciones e interpretaciones de las mismas.

Una vez analizando el informe por sus destinatarios, debe ser negociado con ellos, siguiendo los acuerdos iniciales y estableciendo un análisis compartido que abra paso a la toma de decisiones y a la mejora de de la profesionalidad. Este mismo análisis se convertirá en una forma de autocrítica y de revisión de suma utilidad para el cambio.

El evaluador ha de ayudar a los miembros de la comunidad (o los patrocinadores del trabajo) a evitar algunos posibles errores en la valoración del informe:

a) Errores de tipo A

Aunque las conclusiones del análisis sean rigurosas, NO se aceptan porque resultan molestas. Se escuchará la postura en excusas del tipo “el evaluador no es competente”, “la evaluación ha tenido problemas”, “las conclusiones no son significativas”, “nadie de fuera va a saber mejor que nosotros lo que pasa en el Centro”, etc.

b) Errores de tipo B

Aunque las conclusiones del análisis NO sean rigurosas, se aceptan si los resultados son gratificantes.

En este caso, se disculparán las deficiencias del método subrayando la importancia de los resultados: “lo que se dice en la evolución es muy preciso”, “estaba muy claro desde el principio”. “la evaluación externa es muy objetiva”, etc.

El carácter técnico de la evaluación servirá de apoyo al patrocinador y a la comunidad para utilizar las conclusiones. Es fácil apoyarse en la técnica y en los técnicos para potenciar las tesis y las decisiones.

Cuando existe un patrocinador, la evolución se convierte en un puente de información entre él y los participantes de la experiencia evaluada.

Con ambos tiene obligaciones el evaluador al redactar y entregar el informe:

Con el patrocinador: Salvando su independencia psicológica, social y política, ha de expresar con claridad lo que ha encontrado en su proceso evaluador, respetando los plazos previstos y explicando los desajustes temporales si es que han existido.

Con los participantes: Manteniendo el anonimato de las informaciones recibidas, si así lo ha garantizado y ofreciendo honesta y rigurosamente los resultados obtenidos, independientemente de lo gratificante o molesto que pueda ser el contenido.

La evaluación no es un proceso que se cierra en sí mismo. Tiempo por finalidad mejorar la realidad. Y, aunque no es tarea del evaluador llevar a la práctica los resultados del informe, sí es necesario que facilite las cosas para que así suceda. La evaluación supone un juicio sobre la realidad que se proyecta sobre su futuro. No es una simple mirada retrospectiva.

“la evaluación siempre implica un juicio. Como toda acción expresa una atención y la intención de una evaluación es mirar al paso con la vista en el futuro”. (Hot. 1981).

Uno de los indicadores de la bondad de una evaluación es la voluntad que despierta en la comunidad escolar de iniciar futuras iniciativas de este tipo. Es preocupante que muchas experiencias de evaluación sean ÚNICAS en su género, a causa de su forma de llevarse a cabo y de concluirse. Si la evaluación realizada genera reticencias, rechaza y hostilidad frente a posteriores y posibles experiencias, habremos de concluir que este efecto secundario ha sido un grave inconveniente que no podemos ignorar al hacer la meta evaluación.

El informe ha de ser negociado con la comunidad escolar o con sus representantes, de forma que se convierta en un elemento de mayores virtualidades educativas, al revertir el conocimiento aportado sobre la toma de decisiones. Esta negociación confiere al informe una mayor credibilidad si el evaluador tiene en cuenta las indicaciones de los protagonistas.

No pueden olvidarse los elementos éticos del proceso evaluador (Adelman, 1980; Cook, 1980). Muchos factores relativos a la vertiente política y ética de la misma se hallan vinculados al informe y a su utilización por los evaluadores, por los patrocinadores y por los usuarios.

EVALUACIÓN EXTERNA VERSUS AUTOEVALUACIÓN DEL CENTRO ESCOLAR.

Evaluación externa.

Se considera necesaria la evaluación externa del Centro, ya que pueden realizarla expertos con mayor independencia que los implicados directamente en la acción, con más tiempo dedicado a la tarea y con más rigor dada su especialización técnica. La

autoevaluación encierra el riesgo de que cada grupo actúe con los criterios que justifican aquello que están haciendo o que apoyan la eliminación de lo que resulta molesto.

Ahora bien, si el Centro no acepta esa actividad será difícil que tenga un resultado positivo la evaluación: se obstaculizará el desarrollo y se ignorarán los resultados.

Si la iniciativa procede del poder, es fácil que la evaluación se convierta, a los ojos de los miembros de la comunidad escolar, en un mecanismo de control, ya que en el poder se hayan arraigados componentes políticos, económicos e ideológicos. La evaluación brindará a los patrocinadores oficiales la ocasión de apoyarse en el poder tecnológico. Puede olvidar el resultado de la evaluación si así le interesa y apoyarse en él cuando desea tomar una decisión de que le conviene.

“*Esta es la decisión que aconsejan los técnicos*”, dirán.

El Centro puede considerar al evaluador como una persona en convivencia con el poder. De ahí se defiende de él. O bien, que trate de ganarle para su causa, brindándole una imagen mejorada de la realidad escolar.

De ahí la importancia de establecer estrategias de negociación iniciales que planteen la evaluación como una experiencia enriquecedora para el Centro, como una oferta interesante de la que se pueden extraer beneficios profesionales, desmontando la idea de que es un fenómeno inquietante y amenazador. De ahí, también, la necesidad de dar fuerza a los evaluadores, de darles más poder. En teoría al menos, tienen menos intereses comprometidos en el resultado de la evaluación. Menos que el poder, por una parte, y menos que los integrantes de la comunidad, por otra.

Una de las formas de dar poder a los evaluadores es producir conocimiento sobre la política o ideología de la evaluación. Tener conocimientos sobre la evaluación institucional es una forma de defenderse de los políticos y de los mismos evaluadores. Si los profesores tienen un conocimiento suficiente del proceso evaluador, podrán decir ante un informe cuyos resultados no comparten: “*No lo creo. Yo tengo otros datos*”.

Los evaluadores pueden *presionar* al poder para que actúe teniendo en cuenta las descripciones y las teorías adecuadas. Si no lo hace así, no le pueden obligar a obrar de una forma más honesta y coherente, pero sí pueden plantear esta amenaza: “*Se harán públicas vuestras intenciones*”.

La tarea del evaluador es hacer posible la crítica institucional. Cuando el patrocinador quiere encauzar y condicionar la evaluación, cuando pretende detenerla si continúa por una vertiente determinada que no es de su agrado, al evaluador sólo le queda la posibilidad de consignar esa actitud: “*Esa actitud que también se evaluará*”.

Es imprescindible defenderse de la fiscalización, crear condiciones que garanticen la independencia de los evaluadores en todas las fases del proceso: elaboración del proyecto, desarrollo y elaboración del informe.

La tecnología (la ciencia) se alía fácilmente con el poder. O mejor: es fácil que el poder utilice la tecnología a su capricho, conforme a sus intereses. El evaluador ha de ser conciente de que tiene un poder técnico que puede ser utilizado políticamente por el patrocinador.

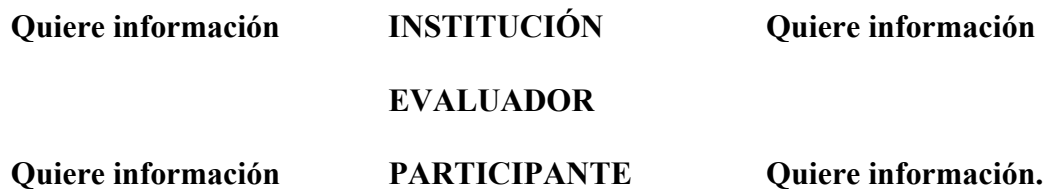
¿Cómo aminorar los riesgos de esta falsificación, de esta utilización abusiva?

- Haciendo más participativo el proceso después de la evaluación. Si el Centro conoce los resultados, será difícil que el poder opere utilizando de forma sesgada los informes.

- Potenciando la fuerza del equipo evaluador, de tal modo que sea factible una negociación justa y equilibrada con el poder.
- Descubriendo y haciendo públicos los mecanismos por los que se introduce la trampa y el engaño.
- Desarrollando el conocimiento sobre la técnica evaluadora. Si hay conocimiento sobre la evaluación (no mitificación cultivada a través de un lenguaje críptico) será más difícil el abuso del poder.
- Evitando las falsas dicotomizaciones: conocimiento científico versus conocimiento vulgar, conocimiento experimental versus conocimiento cualitativo, dando a los otros un valor casi mágico.

Hay que crear estrategias para negociar con éxito (en un sentido y en otro; con el poder y con la comunidad escolar) en situaciones cerradas, pero desde una postura de fuerza que esté presidida por la independencia, la ética y el rigor.

El evaluador es un puente de información entre el patrocinador y los participantes.



Es probable que existan presiones, desde una y otra parte, por conseguir información rápida. Esa dinámica de la presión hace que se produzcan diversos problemas:

- Se exige información en momentos y formas inoportunos.
- Se condiciona al observador para que ofrezca información con un sentido determinado,
- Se clasifica la descripción de la realidad para que el evaluador tenga una visión deformada (que, a su vez, difundirá a las instancias pertinentes).
- Se utiliza la información recibida del evaluador de forma partidista, arbitraria e interesada.
- Se pretende dirigir la evaluación para que llegue a los lugares, personas, problemas y situaciones que interesan.

El papel del evaluador puede resultar incómodo si hace ver las diferencias o los errores, pero su imagen quedará deteriorada si se convierte en un notario de la realidad del Centro, sin capacidad de análisis y de crítica.

La evaluación externa puede ser solicitada por el propio Centro, que es a la vez, entonces, patrocinador y participante, en este caso, es el propio Centro el que debe aplicar los planteamientos éticos y funcionales que antes dirigíamos al poder:

- Al elegir el equipo que realice la evaluación, el Centro está haciendo una opción importante. Si desea conseguir un informe conformista y laudatorio se dirigirá a equipos de “línea”, plenamente identificados con la filosofía del Centro. Si el equipo está “comprado”, hará y dirá lo que el Centro sugiera. De ahí la necesidad de elegir un equipo en el que esté garantizada la competencia científica y la libertad de acción. Mientras más “fuerte” sea el equipo más garantías tendrá de independencia.

- Aunque resulta lógico que si el centro ha solicitado y puesto en marcha una evaluación, facilitará los cauces para desarrollarla y para que se conozcan y difundan los resultados, puede frenar esta disposición por el hecho de que la evaluación discorra por cauces no previstos y/o deseados.
- La utilización de los informes deberá ser democrática y racional, propiciando debates serios y rigurosos que faciliten el ejercicio reflexivo y autocrático y la toma de decisiones coherentes.

Puede también realizarse una evaluación externa en un Centro a petición de una persona o equipo que desea realizar una investigación o emprender una actividad de prácticas. En este caso, hay que asegurar las condiciones para que el ejercicio no sea baldío para el Centro. Para ello hay que asegurar las condiciones que faciliten un desarrollo serio y riguroso y una imprescindible materialización del informe, entregado a la comunidad escolar para su estudio. Muchas de estas investigaciones se cierran cuando el evaluador abandona el Centro, sin dejar más que un rastro de incomodidades y de perturbaciones en la vida ordinaria del Centro.

Existe también el peligro de que la persona que ha sido recibida generosamente por el Centro no tenga el valor suficiente para entregar un informe en el que aparecen interpretaciones cargadas de valoraciones negativas. Es preferible no entregar nada al Centro que distorsionar los resultados en aras de una pretendida y cortés gratitud.

El evaluador externo de un Centro escolar tiene que ser, a mi juicio, un experto en educación. El antropólogo, el sociólogo, el psicólogo, el filósofo, pueden aportar planteamientos metodológicos y visiones teóricas que faciliten la comprensión de la *cultura* del Centro y de las aulas, pero las claves de interpretación más precisas surgirán del conocimiento del fenómeno educativo y de sus especialísimas características.

La evaluación externa debe realizarse en equipo: tiene más rigor el contraste, más fuerza social la tarea, más apoyo psicológico la relación, mayor resistencia a las presiones, más posibilidad de completad el trabajo...

El equipo ha de ser homogéneo en cuanto a la línea teórica ya que hará viable un enfoque coherente y una interpretación de los datos comprensible y uniforme, aunque plural.

En España no existen muchos equipos preparados que puedan realizar estas tareas ya que, hasta ahora, ni las Facultades ni otros Centros de formación desarrollan currícula pertinentes para realizar estas funciones, ni existe una demanda social de estos especialistas por parte de las instituciones.

El funcionamiento de estos equipos profesionales, daría relevancia a la tarea y la iría prestigiando, ya que la experiencia, la preparación científica y los casi inevitables resultados en la mejora de la actividad profesional de los evaluadores harían crecer la sensibilidad y la consideración intelectual hacia el proceso evaluador.

Bien es cierto que experiencia no se puede identificar con ciencia. Por eso será preciso conseguir situaciones que propicien una actividad reposada, rigurosa y libre. Una de las formas de asfixiar una realidad naciente es hacer que las primeras experiencias sean desvalorizadas.

La evaluación externa debería completarse con la autoevaluación institucional. En realidad, la evaluación externa, cuando no ha sido fruto de una iniciativa del Centro, debería desembocar en la puesta en marcha de mecanismos de autoevaluación institucional sistemática.

A mi juicio, la solución se encuentra en la coevaluación, en la combinación de la evaluación externa e interna.

Autoevaluación institucional

La necesidad de la autoevaluación institucional radica en su capacidad para el desarrollo curricular y para la mejora de la profesionalidad de los docentes.

“Las raíces de la reciente importancia de la autoevaluación pueden encontrarse en cuatro vertientes del desarrollo: accountability; desarrollo curricular; revisión curricular y desarrollo del staff” (Clift, Nuttall y McCormick, 1987).

La autoevaluación institucional es una necesidad no suficientemente sentida y ejercida. Cada profesor actúa en su aula y se preocupa, fundamentalmente, de la evaluación de sus alumnos.

No se concibe el trabajo colectivo del Centro como una estructura dinámica en funcionamiento. De ahí la ausencia de una pregunta, por otra parte tan lógica: *¿Qué está sucediendo aquí?*

La lógica exige el análisis del valor del programa de acción que se desarrolla en el Centro, de los lazos existentes entre los diferentes procesos que se dan en la escuela y en la ósmosis de la escuela con la sociedad.

La autoevaluación, como un proceso por el que la institución se mira a sí misma, tiene en su raíz una carga de coherencia importante: *¿Estamos haciendo lo que nos proponemos? ¿Estamos consiguiendo lo que buscamos? ¿Estamos consiguiendo precisamente lo contrario? ¿Por qué?*

“Una de las premisas más importantes en las que se basa la autoevaluación en las escuelas es que los maestros y la escuela son siempre autoevaluables y tienen la preparación y el conocimiento para describir y analizar sus políticas y prácticas en un contexto relevante. De cualquier manera, no es todavía una práctica común (aunque está creciendo) y es necesaria una cierta cantidad de tiempo para que las escuelas produzcan autoinformes y para que los maestros y la gentes de fuera se familiaricen con el tipo de datos ofrecidos y con el criterio con el que deberían ser valorados” (Simona, 1881).

La autoevaluación puede ser total (de todos los aspectos del Centro) o centrarse en algunas parcelas concretas, según las experiencias temáticas, la disponibilidad de tiempo y la urgencia en la solución de problemas. Puede analizar resultados de estudiantes, cómo se organizan los horarios, cómo son las relaciones, cómo se distribuye el espacio, cómo se trabaja con distintos métodos, etc. aún en el caso de que interese el análisis de algún aspecto parcial, será preciso tener en cuenta la compleja estructura y dinámica de todo el Centro. No es posible encontrar significado de una parte sin tener en cuenta el lugar que ocupa en la estructura funcional.

No se puede olvidar que los profesores no disponen de un tiempo para esta tarea. Y aquí encontramos ya una compilación que debe solucionarse a través de una disponibilidad temporal para tareas de tan decisiva importancia. Hay que buscar formas y tiempos de discusión sobre el funcionamiento de la escuela. Es la manera de mejorar la enseñanza y el currículum. No es bueno que el profesor tenga que recurrir al voluntarismo de quien

entrega horas de su tiempo libre para realizar a una actividad que nace de la exigencia misma del funcionamiento institucional.

La autoevaluación encierra el peligro de que la falta de distancia emocional y el compromiso de los agentes conviertan el proceso evaluador en un factor de auto justificación:

“El gran temor de la autoevaluación se convierta en una auto justificación más que en una autovaloración es la que sirve para reforzar la mediocridad más que para promover el cambio necesario. El peligro es quizás más grande en las pequeñas escuelas”. (Nittall, 1981).

Existen problemas que frenan los proyectos de autoevaluación y que deberían ser atajados para que no se viese rota y/o amenazada en su misma sustancia:

- **La resistencia de algunos profesores a ser observados, a ser objeto de evaluación.**

Hay profesores extremadamente celosos de la intimidad de su parcela. Se resisten a ser observados, a ser entrevistados, a ser objeto de la mirada del evaluador. En parte porque están inseguros, en parte porque están escarmentados.

Algunos profesores no pueden pensar que *alguien* pretende apoyarse en la evaluación para perjudicarles (expulsarles de un Centro privado, frenar su promoción, excluirles de un proyecto, etc.) o para ridiculizarles.

No conviene forzar a nadie a ser objeto de la evaluación. De hacerlo, se perdería, a priori, toda la eficacia de la misma en su potencialidad de transformación y de cambio.

- **El carácter individualista de la función docente.**

En nuestro país, donde la preparación inicial y el ejercicio profesional docente se asienten sobre el derecho de la libertad de cátedra y en la soledad compartida de sus cuatro paredes.

- **La falta de motivación profesional.**

Los profesores suelen empezar con entusiasmo, pero son paulatinamente frenados por situaciones frustrantes. (Santos Guerra, 1979; Esteve Zarazaga, 1988). Es difícil que un profesor poco motivado emprenda una tarea compleja y difícil, sin otro aliciente que la mejora profesional.

La ironía de los profesores menos comprometidos, el rechazo y la falta de colaboración de algunas personas, las acusaciones que descalifican su tarea, las malas condiciones institucionales, etc. son obstáculos en los que tropiezan muchos docentes cuando plantean en los Centros la necesidad de realizar una evaluación institucional.

- **La carencia de tiempo.**

Las exigencias de tiempo de la evaluación que aquí plateamos son considerables, aunque se pretenda evaluar una parcela minúscula de la vida del Centro. ¿Cuándo se realizan las entrevistas, cuándo se hace la observación se redactan los informes etc.?+

El profesor, que tiene contabilizadas sus horas en régimen de dedicación de estricto cumplimiento laboral, no dispone de horas libres para este menester. Máxime cuando se trata de una tarea que debería estar complementada como una exigencia de la propia actividad institucional.

“Los impedimentos de la autoevaluación de la escuelas citados frecuentemente por los profesores son tiempo y motivación. La mayoría de los profesores piensan que no tienen el suficiente tiempo para comprometerse con esta “actividad” adicional., y muy pocos son los que están preparados para darle prioridad, especialmente cuando, como pasa a menudo, no hay compromiso firme por parte de la jerarquía escolar de considerar los resultados” (Simona, 1981).

- **Falta de credibilidad.**

¿Cómo dar credibilidad a la tarea de autoevaluación que pretende realizar un equipo de profesores del Centro?

Esta credibilidad es requerida por el mismo equipo de profesores que impulsa y realiza la evaluación, pero también por el resto de los compañeros.

El “amigo crítico” (un profesor de otra escuela, el técnico externo...), el “informe minucioso”, la lectura especializada, el contraste con otras experiencias, etc. son medios de contrastar la actividad autoevaluada.

- **Retraso del momento preciso.**

Existe propensión a retrasar el momento de comenzar, ya que nunca se dan las *condiciones ideales*. En *otro momento, en otras condiciones, con otras personas, con otros medios*, sería posible e incluso deseable. *Pero ahora mismo, no.*

El verdadero problema no es el momento, sino la voluntad y la decisión de llevar a cabo una experiencia autoevaluada, y, en última instancia, el disponer de unas condiciones mínimas que garanticen el desarrollo de la misma.

- **Ocultación de los problemas sustantivos.**

Al ser concebida la enseñanza como un fenómeno individualista, los problemas generales de carácter fundamental se ocultan, se aminoran o se desvirtúan.

Cada uno afronta su problema, su actividad, haciéndose culpable de los errores y limitaciones y acreedor de todos los éxitos, instalándose en explicaciones poco rigurosas, tendentes a mantener la estructura básica. Así, el débil tiene miedo y el fuerte pretende solucionarlo todo por su cuenta.

La estructura funcional sigue operando con los mismos vicios o con los mismos aciertos casualmente hallados y rutinarios mantenidos.

- **Impaciencia por obtener resultados.**

Los profesores suelen tener excesiva prisa por obtener información detallada y resultados inmediatos. Quieren saber pronto lo que sucede, a qué causas se debe, quiénes son los responsables...

Como el proceso de evaluación etnográfica es lento y poco terminante (las evaluaciones de carácter cuantitativo suelen ofrecer datos más “concluyentes”) los profesores se impacientan y tienden a dejar de lado una tarea que tiene frases poco especulares en cuanto a los resultados y logros visibles.

Una de las ventajas de la autoevaluación es que permite poner en marcha procesos de autoanálisis. La autoevaluación no se fija tanto en los logros cuanto en el potencial de reflexión: ¿qué es lo que dinamiza la acción?, ¿qué hay de nuevo en ella?, ¿Qué cambios claves se precisan?, ¿qué limitaciones se encuentran?, ¿Por qué suceden precisamente estas cosas?, ¿Qué nuevas perspectivas aparecen en el horizonte?...

La preocupación por los logros, además de revestir una gran complejidad en cuanto a la naturaleza de los mismos, esconde otros problemas de gran envergadura: ¿a qué precio?, ¿con que plazo?, ¿a qué costa?... Un planteamiento autoritario en un Centro/aula puede producir efectos inmediatos de eficacia en la acción y orden establecido, pero puede producir inseguridad y desconcierto a largo plazo...

Hay logros que se producen “al margen” e incluso “a pensar” de la organización. Y está suele ser muy hábil y pronta para atribuírselos. (La responsabilidad de los fracasos suele atribuirse a la mala respuesta de los individuos). Hay *antilogros* (efectos negativos laterales, no pretendidos) por los que casi nunca se pregunta la institución.

El evaluador debe proteger lo que las personas intenten crear. Hay que observar y aprender muchas cosas sobre problemas, pero reservar la crítica descalificadota. El evaluador puede decir: “*Naturalmente, hay errores*”. Pero está salvando la voluntad de los individuos.

Debe existir también un momento para decir: “*Quizás sea ésta una buena idea aquí. No en otro sitio, no en otra situación, no en otro momento. Pero aquí, si*”.

Hay que construir la curiosidad, la apertura a la búsqueda permanente. No hay que crear modelos perfectos que acaban siendo asfixiantes.

Es preciso conseguir en la autoevaluación que los profesores entiendan lo que está haciendo para que se reduzca el temor y se ajusten las expectativas.

Hay que conseguir que los padres entiendan la autoevaluación del Centro como un medio de mejorar la actividad y no como un ajuste de cuentas al profesorado y un modo de exigir que cumplan con sus obligaciones y responsabilidades.

Hay que procurar que la autoevaluación no se convierta en una simple operación de cosmética, que no cambia nada esencial y se constituye en un adorno sin trascendencia.

La autoevaluación exige, para ser efectiva, que se rompan las barreras que pueden levantarse por la jerarquía, la territorialización y la privacidad (Simona, 1986). De lo contrario, los núcleos de conflicto y de estancamiento permanecerán intactos.

A. Jerarquía.

La jerarquía debe abrir las puertas de su reducto para que los evaluadores puedan actuar con libertad. Si la autoridad actúa preservando con celo su reducto, no será posible analizar

con claridad los procesos que se desarrollan en el Centro y tenderán a multiplicarse las líneas de resistencia.

También es cierto que la descripción de esos mecanismos de protección jerárquica, explican un modo de comportamiento significativo. Si algunos dicen en el Centro: “No se puede hablar sobre el Director”, ya se ha comentado con elocuencia mucho sobre su forma de actuar.

B. Territorialización.

Hay muchos territorios con fronteras demarcadas, por ejemplo, un Seminario, el Consejo Escolar, la Sala de Profesores... Estos cotos pueden cerrar sus puertas a la presencia de observadores que realizan la evaluación.

Estas fronteras deben caer en aras de la transparencia, de la ejemplaridad y de la coherencia.

No siempre es posible, porque hay profesores que ofrecen resistencia y racionalizan la obstaculación en aras de su autonomía y de la eficiencia: “Se dificulta la vida del Seminario”, “Nosotros sabemos cómo funciona esto...”, “esto es una pérdida de tiempo”, “La evaluación no resuelva nada”..., etc.

C. Privacidad.

El profesor se siente dueño de su aula. No siempre está dispuesto a permitir que un observador altere la actividad normal que realiza con los alumnos.

En este caso, como en los anteriores, no es bueno derribar por la fuerza las fronteras, invocando el interés común. Es probable que, con esa violencia, sólo se consiguiesen efectos negativos.

La mejor forma de promover la autoevaluación es que un equipo de profesores ponga en marcha este proceso y se hagan patentes o beneficios de la misma para la mejora de la actividad y para el funcionamiento más eficaz y satisfactorio.

Yo añadiría otra condición para que fuese efectiva la evaluación de la escuela. Es la autonomía de profesores y del Centro mismo. Si no existe autonomía real, de poco sirve la evaluación y prácticamente nada puede conseguirse desde la dinámica interna del cambio. “Las escuelas y los maestros, dice Hoyle (1986b) necesitan un grado de autonomía si quieren ser eficaces”.

PINEDA, Moctezuma, Angélica. “¿Qué es la evaluación del aprendizaje?” En: Evaluación del aprendizaje (guía para instrucciones). México. Ed., Trillas, Pp. 19-28.

2. ¿Qué es la evaluación del aprendizaje?

Al evaluar el proceso de capacitación se puede analizar la forma en que se hace:

- a) La *detección* de necesidades,
- b) la *planeación* didáctica,
- c) el *desarrollo* del evento,
- d) la *aplicación* de lo aprendido en el trabajo real.

Es conveniente que el análisis sobre las necesidades y la forma en que se ha señalado un evento se efectúa *previamente* a su realización, aun cuando algunos elementos se puedan retomar durante las acciones de capacitación. La evaluación de los resultados “reales” de la capacitación, idealmente debería efectuarse después de terminados los cursos, ya fuera en el propio trabajo o en acciones de capacitación subsecuentes. Sin embargo se debe reconocer que, a pesar de su importancia, pocas veces se efectúa.

La evaluación en el desarrollo de los eventos puede referirse: a los resultados inmediatos que se obtienen, en términos de aprendizaje; a la forma en que se llevó a cabo la instrucción: a los elementos materiales con que se contó, etc. Si bien este análisis se puede hacer después de concluidas las acciones, es más conveniente que se practique en el mismo evento.

Repasemos:

En el desarrollo de un evento de capacitación (curso, foro, seminario, etc.) se debe evaluar:

- a) si las acciones se realizan de acuerdo con lo programado, detectando las desviaciones para poderlas corregir,
- b) si los medios, métodos, técnicas, materiales didácticos y, en general., todos los recursos, contribuyen o no al logro de objetivos, y
- c) si se alcanzan los objetivos propuestos, en qué medida, ubicando los resultados inmediatos de las acciones de capacitación, en términos de aprendizaje.

en esta guía sólo se hace referencia al *tipo de evaluación que se hace durante los eventos, para estudiar el aprendizaje que se alcanza en ellos.*

2.1 APRENDIZAJE

Es necesario hacer algunas reflexiones sobre el concepto de aprendizaje antes de seguir adelante:

Aprender no es únicamente almacenar información, sino que *es un proceso global que abarca a la persona en su conjunto y que la lleva a actuar, a percibir y a comprender las cosas de distinta manera.*

En esta guía se concibe al *aprendizaje* como un *proceso dinámico de modificación de pautas de conducta que está en continuo movimiento, involucrando a la persona en su totalidad. Afecta no sólo los pensamientos de un individuo, sino el conjunto de sentimientos y emociones con los que actúa.*

Hay que tener cuidado: “modificación de pautas de conducta” no es lo mismo que “modificación de conductas”, conforme se entiende en un enfoque “conductista”. Veamos a qué se hace referencia al hablar de pautas de conducta:

Si pudiéramos observar todo o que hace un individuo a lo largo de un día, una semana, un mes, y si, además de observarlo, pudiéramos platicar profundamente con él sobre cómo vive las experiencias, podríamos ver cómo muchas de sus conductas (que a primera vista podrían parecer independientes y desligadas unas de otras) son emitidas siguiendo un mismo modelos o una misma

pauta... Cada conducta nueva es emitida sobre la base de otras anteriores; cada conducta nueva es estructurada en forma semejante a conductas pasadas; es decir, que en situaciones parecidas a otras anteriores, ante objetos semejantes a otros del pasado (o que el sujeto capta como semejantes), la persona reacciona de una manera parecida a como lo hizo en la ocasión anterior.

Una pauta de conducta es, pues, una determinada forma de estructurar la conducta, que se repite de tal manera que llega a constituir parte integrante del sujeto. Estas pautas de conducta son modos privilegiados de comportamiento que el sujeto ha seleccionado (no siempre conscientemente) de entre todos los modos posibles, y que en su conjunto caracterizan la personalidad de dicho sujeto.⁸⁴

Por eso, si se considera el aprendizaje como la modificación de pautas de conducta, se habla del cambio general de comportamiento, que permite al sujeto actuar de manera más adecuada en nuevas situaciones, aun cuando su tendencia sea la de seguir actuando en la manera que ya conoce y que puede ser buena o mala, pero que “es la suya”.

Una persona tiene una serie de esquemas cognoscitivo-afectivos con los que explica el mundo y se ubica en él. Al aprender, esos esquemas se confrontan con nuevas percepciones, se modifican o reestructuran, integrándose nuevos “esquemas referenciales”, desde los cuales, con una visión diferente, se vuelve a aprender, los esquemas referenciales modificados llevan al individuo a efectuar algunos cambios en su forma de vida (opiniones, habilidades, etc.), de actuar, de sentir o de entender el mundo.

Al ser un proceso de cambio, el aprendizaje se va a encontrar con ciertas reacciones de resistencia y mecanismos de defensa, conscientes e inconscientes, y atravesará una serie de pasos.

Los pasos o fases que atraviesa son:⁸⁵ *apertura, desarrollo y culminación.*

Al recibir una nueva información (sea una plática, la observación de un acontecimiento, etc.) ésta se percibe como totalidad, con cierto grado de indiscriminación. A través de un esfuerzo mental se analiza, descomponiendo la totalidad a partir del apoyo en elementos teóricos explicativos, para poder construir nuevas síntesis que permiten la elaboración de nuevos cuestionarios y posibles explicaciones, que llevan a reiniciar un proceso de “destotalización” en análisis posteriores.

En el aprendizaje se puede avanzar o retroceder, dar saltos y cambios profundos, o paralizarse por miedo o porque esquemas anteriores, con repercusión afectiva, se ven amenazados. Esto se conoce como *espiral dialéctica del aprendizaje.*

El aprendizaje es un proceso social. Cuando el esquema referencial individual entra en contacto con el esquema referencial del otro, se posibilita la apropiación y transformación del conocimiento de un proceso de construcción y movilización del esquema referencial propio.

El aprendizaje no se puede considerar como una asimilación o absorción pasiva, sino el aprendizaje es estrictamente individual.

Aunque, al igual que se decía del proceso de conocimiento, puede darse en cualquier situación (espontáneo). Constituye la meta del proceso de enseñanza-aprendizaje (intencionado). En este último, aprender conjuntamente instrucciones y capacitandos. No

⁸⁴ Carlos Zarzar Charur, “conducta y aprendizaje: una aproximación teórica, en: *Grupos de aprendizaje*, Págs. 36-37.

⁸⁵ Existen varias explicaciones sobre las etapas que atraviesa el aprendizaje. aquí nos basamos en la propuesta por Azucena Rodríguez. Estos tres momentos de aprendizaje corresponden a las fases de Asimilación y Acomodación de que habían otros autores (Piaget, Hilda Taba).

sólo corresponde a los participantes aprender. El instructor cambia a su vez, como resultado de su interacción con los capacitandos. Es un proceso educativo del grupo, en que todos sus integrantes interactúan aprendiendo.

Síntesis:

*El aprendizaje se concibe como un proceso de **modificación de pautas de conducta**, por medio del cual se adquieren, retienen y utilizan conocimientos, habilidades, actitudes, hábitos e ideales, produciendo cambios tanto simples como complejos en las respuestas, reacciones, pensamientos, actitudes y, en general el comportamiento del que aprende.*

2.2 PROCESO DE ENSEÑANZA-APRENDIZAJE

La enseñanza y el aprendizaje son procesos que están íntimamente ligados, es decir, son interdependientes. La enseñanza se concibe como un proceso intencional donde interactúan maestro alumno y del que el aprendizaje buscado es el principal resultado.

Enseñar no es transmitir conocimientos; no se pueda dar aprendizaje. Sin embargo, el aprendizaje es la razón de la enseñanza.

*La **enseñanza** consiste en la dirección de actividades y la utilización de elementos para facilitar a una persona, por su propia actividad, aprender.*

Algunos autores prefieren referirse a la idea de “propiciar situaciones de aprendizaje”, que a los conceptos de enseñanza o instrucción.

Por la estrecha interacción mencionada, se considera que al evaluar el aprendizaje no sólo se valora a los participantes e instructores, como grupo que aprende, y los resultados de su interacción: el aprendizaje.

El primer elemento que se debe considerar en cualquier proceso de aprendizaje intencionado y estructurado en la delimitación de objetivos. Éstos constituyen las metas por alcanzar, permiten establecer la dirección del proceso y planificar las experiencias y situaciones de aprendizaje. Es conveniente o que existe un acuerdo con los participantes, acerca de los objetivos que conjuntamente procurarán alcanzar.

2.3 EVALUACIÓN DEL APRENDIZAJE

Pero volvamos al tema de esta guía. Existe una serie de ideas en tomo de evaluación del aprendizaje; que “evaluar los conocimientos” consiste en saber “cuánto se aprendió”, a través de una serie de preguntas sobre la información sea idéntica a lo expuesto por el instructor, será correcta; que se puede calificar por un número –o letra-; y que este número representa lo que una persona “ha aprendido”. Estas ideas suponen una visión mecánica y estática del aprendizaje.

Si, como proponemos, aceptamos que el aprendizaje es un proceso dinámico, en donde quien aprende debe modificar conscientemente sus propios esquemas cognoscitivos, esto nos lleva a una visión diferente de la evaluación del aprendizaje. Su objeto de estudio es el proceso de aprendizaje, su objetivo es conocer el proceso y resultados del proceso de aprendizaje.

En la capacitación, la evaluación del aprendizaje es un proceso sistemático de investigación, que realizan conjuntamente capacitandos e instructores para valorar, en el transcurso de la capacitación, el desarrollo de los conocimientos, habilidades y actitudes propuestos, con el fin de retroalimentar el aprendizaje de los participantes, la actuación de los instructores y el proceso mismo de capacitación.

La evolución del aprendizaje tiene, necesariamente, un carácter limitado y parcial, ya que por el dinamismo y complejidad de su objeto de estudio (el aprendizaje), solamente obtiene “cortes” –artificiales- en este proceso, es decir, *sus resultados representan únicamente lo que se aprendió acerca de algunos aspectos, hasta el momento en que se hizo la evaluación.*

Sin embargo, estas limitaciones no demeritan su práctica; la aplicación de la evaluación producen por sí misma una modificación en los esquemas cognoscitivos; contribuye a que los participantes tomen conciencia de sí mismos, de su situación; a que reflexionen sobre su aprendizaje (para qué aprendieron y cómo lograron aprender) al igual que posibilita que se analice la relación entre las diferentes personas que intervienen en el evento, contribuyendo a modificar las concepciones que sobre el aprendizaje tienen los capacitandos y los instructores.

Como ya se mencionó en el capítulo 1, la capacitación se caracteriza por ser una acción intencionada. Quienes la patrocinan (la organización, la empresa, la institución, etc.) buscan, que quienes participan en un evento alcancen cierto aprendizaje sobre su poder la evaluación:

1. Evaluación como comprensión del proceso de aprendizaje, en el grupo y en los individuos, en la que el principal propósito es promover la reflexión y la participación en el análisis colectivo.
2. Acreditación, que se relaciona con el requerimiento institucional de certificar si se obtienen los conocimientos que se marcan en los objetivos. Sólo en este caso es preciso emitir una calificación.

Recordemos que cualquier evaluación parte de un sistema de valores. En la evaluación del aprendizaje el parámetro lo constituyen los objetivos del evento.

Las apreciaciones sobre el aprendizaje, o resultados del análisis evaluativo deben incluir:

- a) Descripciones cualitativas de lo aprendido y, cuando sea necesario calificar, descripciones cuantitativas,
- b) Juicios valorativos con respecto a ese aprendizaje (que toman como parámetro los objetivos del evento), y
- c) Propuestas de acción para reafirmar el aprendizaje.

2.3.1 Objetivos de la evaluación del aprendizaje

Estos son:

- a) Valorar el aprovechamiento individual y grupal de los participantes, en cuanto a conocimientos, habilidades y actitudes desarrolladas en el evento,

- b) Retroalimentar constantemente el aprendizaje de los capacitados, señalando las áreas que presentan deficiencias, las que deben corregirse, reafirmarse o ejercitarse,
- c) Proporcionar información que fundamente la reprogramación de contenidos y actividades del curso que se esté desarrollando y de otros posteriores, y
- d) Estimar la eficacia del evento, en cuanto al grado que se alcanzan los objetivos y, con base en éste, la de todo el proceso de capacitación.

2.3.2 Principios de la evaluación del aprendizaje

De acuerdo con lo anterior, la evaluación debe seguir ciertos principios; así pues, debe ser:

- a) *Permanente*, es decir, estar presente en todo el desarrollo del evento.
- b) *Integral*, al evaluar de forma representativa los conocimientos, habilidades y actitudes que se pretenden desarrollar.
- c) *Sistemática*, o sea, proceder conforme a un plan y a un método, con objetivos claros, técnicas e instrumentos adecuados y con criterios de valoración definidos
- d) *Funcional*, al proporcionar información que permita mejorar el desarrollo del evento.
- e) *Didáctica*, es decir, que dé oportunidad de corregir errores y reafirmar conocimientos.
- f) *Participativa*, o sea que propicie la autoevaluación, la reflexión y la crítica de los participantes acerca de su propio aprendizaje.

2.3.2 Requisitos técnicos indispensables

La evaluación del aprendizaje deberá cubrir ciertos requisitos técnicos, a fin de asegurar su calidad y la posibilidad de cumplir con eficiencia sus objetivos.⁸⁶

1. *Validez*. Se refiere al grado en que la información captada por la evaluación y por los resultados obtenidos, constituyen una representación veraz de lo que aprendieron los participantes en el evento, y a través de éste, del logro de los objetivos y expectativas. Es la característica más importante; siempre se debe preferir un procedimiento que evalúe lo que interesa, a otro que aunque sea confiable y práctico, evalúa aspectos sin relevancia.
2. *Confiabilidad*. Se refiere al “grado de consistencia” de la información captada expresa verdaderamente el aprendizaje, y que los resultados obtenidos no deben variar aunque se use una prueba diferente (pero equivalente) para evaluar las mismas destrezas y conocimientos, o aunque califiquen personas distintas.
3. *objetividad*. Se deben eliminar en el mayor grado posible los factores subjetivos que pudieran influir en la elaboración, aplicación y calificación de la evaluación.
4. *Accesibilidad de uso*. Independiente de la eficiencia del procedimiento, se debe cuidar que sea práctico: debe contarse con tiempo y materiales para realizarlo, presentar facilidad para su utilización, registro de información, interpretación de resultados y tener un costo moderado.

EJERCICIO DOS

Reflexione y conteste las siguientes preguntas antes de seguir adelante:

1. La forma en que aprenden los capacitandos con los que usted trabaja, ¿es igual a la forma en que usted aprende? ¿Qué diferencias existen?
(Es decir, habría que preguntarse cómo aprende el indígena, el campesino, los técnicos, etc., como perciben o que se expone, cómo la relacionan con su realidad, etc.)

2. Considerando las características del proceso de aprendizaje, ¿qué opina de que una calificación numérica pueda representar lo aprendido en un evento?

3. ¿De qué forma puede contribuir la evaluación a consolidar el aprendizaje?

4. ¿Qué objetivos y características considera deseables para la evaluación que usted realice?

5. Describa las características de las evaluaciones que ha elaborado en ocasiones anteriores. ¿Qué aciertos y que limitantes encuentra en ellas?

6. Analice sus respuestas a las preguntas 2, 3, y 4 del ejercicio de autoevaluación inicial. Haga las modificaciones que considere necesarias.

3 ¿Cómo se puede evaluar el aprendizaje?

Como se señaló en el primer capítulo, una limitante de la evaluación es la subjetividad, y para superarla es necesario explicitar la estructura de valores de que se parte. En otras palabras, la evaluación debe instrumentarse de *forma sistemática*.

Para ellos se requiere de cierta información previa y del seguimiento de ciertos pasos:

- a) la información previa es la *planeación* (por ciento) del evento, la cual debe incluir, por lo menos:
 - ☞ características de los participantes que se capacitarán,
 - ☞ objetivos generales del evento,
 - ☞ objetivos de cada tema, y
 - ☞ actividades que se realizarán para desarrollar cada tema, con sus tiempos y materiales.

- b) Los pasos que deben seguir para instrumentar la evaluación del aprendizaje:
 - ☞ hacer el *plan de evaluación del evento* en el que se determina qué se debe evaluar, con qué técnica y en qué momento.
 - ☞ *Planear cada una de las técnicas e instrumentos* que se utilizarán mediante la elaboración de una tabla de especificaciones, en la que se relacionan los aspectos por evaluar con sus indicadores y criterios de valoración,
 - ☞ *Construir y probar los instrumentos* que se emplearán en cada una de las técnicas,
 - ☞ *Aplicar* la evaluación en el desarrollo del evento,
 - ☞ *Codificar y calificar* la información que se obtiene de cada técnica (calificar sólo cuando sea necesario), y
 - ☞ *Analizar e interpretar los resultados* obtenidos para formular propuestas de acción para el desarrollo del evento.

Cada uno de estos pasos se explicarán más adelante, después de presentar las técnicas que se pueden utilizar para evaluar el aprendizaje. Únicamente cabe señalar que a todos estos pasos se les puede dar un carácter participativo.

EJERCICIO TRES

Reflexione y realice lo que se indica:

1. Cuando ha participado como instructor en un curso, ¿qué pasos ha seguido para elaborar, aplicar e interpretar los resultados de la evaluación del aprendizaje?, ¿Cómo los ha desarrollado?

2. Elija un curso en el que usted pueda participar como instructor.

Desarrolle por escrito su planeación didáctica, ya sea mediante una carta descriptiva, programa, guías de trabajo, etc. Asegurese de que contiene los elementos mínimos a que se hizo mención en esta apartado.

3. Describa con la mayor precisión las necesidades a que debe dar respuesta la evaluación del evento diseñado en la pregunta anterior. Escriba los objetivos de la evaluación del aprendizaje que usted haría en ese evento (puede retomar lo que elaboró en el ejercicio anterior).

4. Comparte el procedimiento expuesto en este apartado con su respuesta a la pregunta 6 del ejercicio de autoevaluación inicial. Modifique lo que crea conveniente.

4. Técnicas e instrumentos para la evaluación del aprendizaje

La evaluación sistemática del aprendizaje se realiza con el empleo de ciertas técnicas e instrumentos.

Las **técnicas** son el conjunto de procedimientos y actividades que permiten que se manifieste y demuestre el aprendizaje obtenido para poderlo valorar y comparar con los objetivos propuestos.

Los **instrumentos** son los materiales de que se sirve el evaluado para guiar o conducir una técnica y recabar en forma sistematizada la información que se obtiene de ellas.

4.1 CRITERIOS DE SELECCIÓN

Las técnicas e instrumentos deben elegirse en función de:

1. las características de los participantes: escolaridad, alfabetización, experiencia en resolver exámenes, modos de expresión naturales, etc. No se puede evaluar de igual forma a un campesino analfabeto o a un indígena monolingüe, que a un ingeniero agrónomo o a un economista.

2. los objetivos y naturaleza del aprendizaje que se desea evaluar, ya que hay una diferencia entre pretender que se conozca y comprenda cierta información, o que incluso ésta se analice, y pretender desarrollar una habilidad para efectuar una tarea o elaborar un producto.
3. las condiciones con que se cuenta: tiempo, materiales disponibles, ambiente físico, etcétera.
4. la preparación de la persona encargada de coordinar la aplicación de la evaluación.

Cada técnica y cada instrumento debe utilizarse de acuerdo con propósitos definidos, integrarse en lo posible a las actividades de aprendizaje del evento, coincidiendo con los momentos más adecuados –en los que se puede captar lo aprendido de una manera casi natural-, reunir los requisitos técnicos (de validez, confiabilidad, objetividad, etc.) y recabar datos que se articulen y permitan concluir sobre el aprendizaje desarrollado

Resumiendo:

Las técnicas deben adecuarse a las características de los participantes, de los objetivos y del aprendizaje que se desea evaluar, así como a las condiciones con que se cuenta y a la preparación del evaluador.

4.2 PERFIL DE LOS CAPACITANDOS

Los capacitandos son personas adultas que aprender de una forma particular:

- a) Estudia por propia decisión;
- b) Selecciona lo que ha de aprender;
- c) Está dispuesto a esforzarse sólo si percibe la importancia vital de aprendizaje, si lo reconoce como funcional y significativo;
- d) Requiere motivaciones muy concretas y específicas: rendir más en una actividad determinada y de aplicación inminente;
- e) Su concepto de tiempo como “un bien escaso”, le crea un sentido de urgencia;
- f) Su experiencia es amplia; ha adquirido la capacidad de establecer relaciones y de tender múltiples puentes entre los hechos y las personas, y también entre unos hechos y otros;
- g) Es conciente de que el maestro funciona como un simple servidor; su postura es de crítica y exigencia;
- h) Su atención es firme (*en contraposición a la de un niño*) y puede permanecer concentrado largo tiempo;
- i) su aprendizaje es, sobre todo, reorganización, reestructuración, profundización;
- j) por lo general vive atrapado en ideologías y en redes de prejuicios;
- k) desarrollo resistencias y fácilmente se bloquea, por sentirse amenazado ante el cambio.¹

¹ Mauro Rodríguez. *formación de instructores*. Págs. 29-30.

A pesar de que éstas son las características generales, no todas son aplicables en la capacitación dirigida a campesinos, indígenas y técnicos de campo,² por lo que habría que considerar los siguientes aspectos:

a) Pensamiento Concreto. Los individuos o grupos sociales cuyas condiciones de existencia se hallan estrechamente asociadas al trabajo productivo material (trabajadores directos), desarrollan una forma de pensamiento que se vuelve cada vez más concreta, en relación con su proceso material de trabajo, con escaso desarrollo en formas de pensamiento abstracto. El pensamiento concreto se estructura a partir de la formulación y utilización de proposiciones particulares y sobre todo singulares, aun cuando llevan implícitas generalizaciones y proposiciones universales.

El pensamiento concreto tiene una gran riqueza, validez práctica y existencial.

b) Analfabetismo. El no saber leer y escribir contribuye a la marginación y sobreexplotación, a la disminución de sus derechos políticos y procede efectos negativos sobre la autoestima de los analfabetas. Sin embargo, esto no significa que sean ignorantes, ya que cuentan con su experiencia existencial acumulada.

c) Relación medio urbano-medio rural. Económica e históricamente el medio urbano industrial ha preponderado sobre el medio agrícola rural. Aparece una relación en que lo urbano, lo que viene del centro, es lo válido (lo que provoca disgusto, por ejemplo, en los técnicos de campo al recibir disposiciones de oficinas centrales) y se descalifica lo local. En ocasiones, en el indígena o campesino esto ocasiona una admiración o dependencia respecto al técnico (“es el que sabe”) y no valora su propio saber.

d) Visión localista. Como resultado de su inserción en una estructura productiva determinada, la realidad objetiva es vista frecuentemente por el campesino adulto de una manera focal y estática. Las relaciones que determinan a esa realidad y le dan sentido y coherencia, frecuentemente se le escapan.

**SAPEG, Chain Nassir; “Nuevo enfoque para la evaluación de proyectos”
En: Criterios de evaluación de proyectos. México. Ed. McGraw-Hill. Pp. 1-18.**

CAPITULO I

Nuevo enfoque para la evaluación de proyectos

POR QUE EVALUAR LOS PROYECTOS

² Joao Bosco Pinto, “La educación de adultos y el desarrollo rural”, en CREFAL, *siete visiones de la educación de adultos*.

No cabe duda de que hoy en día la preparación y evaluación de proyectos cumple un papel de primera importancia entre los agentes económicos responsables de decidir acerca de la asignación de recursos para implementar iniciativas de inversión.

Frente al tradicional criterio de realizar préstamos para nuevas inversiones en función de las garantías que pudiera ofrecer el solicitante, algunas instituciones financieras se destacaron como pioneras en la incorporación de los análisis profesionalizados de otorgamiento de créditos, generalizando la aplicación de la técnica metodológica denominada de preparación y evaluación de proyectos, para facilitar y mejorar su proceso decisorio.

Con esto lograron no sólo velar por los intereses de la institución otorgante del crédito, sino también por los intereses de inversionista, que, en definitiva, es quien debe responder con sus bienes dados en garantía por cualquier error en la decisión.

Más concretamente, su decisión de optar por una de las posibilidades de endeudamiento disponibles debe tomarse en función del resultado de un análisis que muestre una rentabilidad positiva posible de esperar de una inversión próxima en el tiempo y que busca obtener con ese préstamo.

La realización de un estudio para medir la rentabilidad de un proyecto, sin embargo, no debe verse sólo como un requisito impuesto por una institución financiera para prestar recursos financieros, sino que principalmente como un instrumento que provee una importante información a los inversionistas respecto a su propia conveniencia de llevarlo a cabo.

PRIMERO, FORMULAR

La importancia que se ha dado al estudio de proyecto ha originado que hasta ahora se hayan escrito muchos textos sobre la temática de la evaluación de inversiones. Sin embargo, son muy pocos, o casi ninguno, los que se preocupan de lo que estimamos más determinante de la efectividad de un estudio de esta naturaleza; su etapa previa de formulación y preparación del proyecto.

Quizás el primer intento profundo efectuado en este sentido, y que marcó una línea general que se mantiene hasta hoy, fue el realizado por las Naciones Unidas, organización que publicó, en 1958, el *Manual de proyectos de desarrollo económico*¹, obra de don Julio Melnick. En él se expusieron por primera vez los lineamientos que debían guiar la acción de los evaluadores en la etapa de preparación, como una forma de apoyar la tarea prioritaria de recabar el máximo de información de carácter económico a partir de la cual se evalúan los proyectos.

En esa perspectiva, podría afirmarse que fue este estudio el que dio la pauta para que se pudiera determinar que tanto o más importante que saber aplicar las herramientas financieras de cálculo de la rentabilidad de un proyecto de inversión es saber previamente identificar y cuantificar correctamente todos los costos y beneficios que determinarán en definitiva, su real rentabilidad, así como también la forma como éstos deberán considerarse en dicho proceso.

A pesar de los años transcurridos, recién en 1983 surge, en el departamento de la Administración de la Facultad de Ciencias Económicas y Administrativas de la Universidad de Chile, un texto que incorpora a la estructura general propuesta por

¹ Editado por Naciones Unidas, publicación 5.58.11.G.5., México, 1985

Naciones Unidas una serie de metodologías específicas que proporcionan procedimientos y soluciones técnicas a cada una de las etapas, tanto de la preparación como de la evaluación del proyecto. Asimismo, presenté una base metodológica para la sistematización e integración correcta y efectiva de esas etapas individuales en un esquema estructurado en función de los resultados esperados del estudio de un proyecto².

En 1987, se publica en México un segundo libro³, que, en pocas páginas, resume prácticamente al anterior, aunque sin aportar nuevas soluciones metodológicas. Sin embargo, es una muestra de la percepción de la conveniencia de llegar a un público menos exigente, pero necesitado de información acerca de una técnica que adquiere, día a día, mayor importancia y aceptación entre los agentes económicos que participan en el proceso de una inversión.

UN TEXTO DIFERENTE

Con el nuevo libro que este primer capítulo introduce se busca exponer un ángulo diferente para el estudio de posproyectos, aunque con la misma finalidad de cubrir el área temática más importante, más compleja y, a mi juicio, más frecuentemente evitada en otros textos, como es la determinación de la magnitud de las inversiones, costos y beneficios, o, como generalmente se la denomina, la preparación del proyecto.

Sin embargo, la perspectiva desde la cual pretendemos hacer esta presentación se aleja significativamente de los esquemas tradicionales de un texto de estudio.

Creemos que las bases teórico –conceptuales han sido ya suficientemente tratadas en los textos indicados. Por ello, se ha buscado, a diferencia de los libros editados con anterioridad, más que detenerse en definiciones, fórmulas o procedimientos de cálculo, dar, en general, una clara y completa explicación de los elementos incluidos en un estudio de proyectos y, en particular, llamar la atención sobre los aspectos más determinante del éxito o fracaso del trabajo del evaluador e proyectos.

Fundamentalmente, las páginas que siguen se abocarán a dar la técnica los alcances que realmente tiene, como único mecanismo para poder pedirle y valorar los resultados que ésta en condiciones de otorgar. Junto con ello se tratará, por la vía de múltiples experiencias especialmente seleccionadas, ayudar a comprender la forma correcta de enfrentar un problema de cálculo de la rentabilidad de una posible inversión.

¿MÁS TÉCNICAS O MÁS CRITERIO?

La técnica aplicada sin criterio, las muchas veces excesiva confianza en los resultados del estudio, del desconocimiento de algunas modernas innovaciones metodológicas y la no consideración de la totalidad de las opciones de análisis técnico, sin factores que explican muchos de los errores cometidos por los evaluadores.

² NASSIR SPAG Y REINALDO SPAG, *fundamentos de preparación y evaluación de proyectos*, Santiago: Universidad de Chile, 1983.; reimpresión, Bogotá: McGraw-Hill, 1985. Hay una segunda edición corregida y ampliada, con el título de *preparación y evaluación de proyectos* Bogotá: McGraw-Hill, 1989.

³ GABRIEL BACA, *Evaluación de proyectos*. México: McGraw-Hill, 1987.

En tal sentido, este libro intentará determinar, describir y analizar los errores más frecuentes de las evaluaciones de proyectos, a la vez que proponer algunos procedimientos y criterios particulares para evitarlos y lograr así cumplir con la finalidad última de este estudio: una eficiente asignación de los recursos.

Para estos efectos, en cada uno de los capítulos siguientes se intentará analizar un tema de interés específico, buscando transmitir, más que los modelos conceptuales ya proporcionados en nuestra anterior publicación, los criterios generales de interpretación y particulares de aplicación que posibiliten reducir los errores en la estimación de los costos y beneficios de un proyecto, así como en la forma de medir su rentabilidad.

CUIDADO CON LAS INCOHERENCIAS

Aunque cada capítulo de este texto pretende abordar en forma independiente un tema específico, lo cierto es que el evaluador debe siempre velar por la integración de todas las partes de su trabajo.

Al formular cualquier proyecto de inversión se requerirá generalmente de un trabajo interdisciplinario que obligará al jefe del proyecto a mantener preocupación y una dedicación especial-a la vez que más profunda- por lograr la mayor coherencia de los antecedentes proporcionados por especialistas de muy diversas áreas, para llegar a estructurar un proyecto que mantenga, en la etapa de sus formulación, un equilibrio armónico entre todas sus partes. En el hecho de no hacerlo así radica comúnmente el primer error frecuente de las evaluaciones.

Es asombrosa la cantidad de proyectos que se podrían mencionar, entre muchos que hemos revisado, en los cuales se encuentran incoherencias que implicaron graves efectos sobre el resultado de la evaluación, determinando la mayoría de las veces que el proyecto mostrase una rentabilidad superior a la que debería poder esperarse de un estudio efectuando más eficientemente.

PROBLEMAS CON LAS INVERSIONES

Más frecuentes son, sin embargo, los casos en que no concuerdan las características propias de cada proyecto con el nivel de la inversiones del mismo y, más específicamente, aquella que debe hacerse en capital de trabajo, si está correctamente calculada, probablemente se logre garantizar el funcionamiento normal del proyecto cuando entre en operación.

En una ocasión en que debí efectuar la revisión técnica de un proyecto presentado a una institución financiera, fue notoria la sorpresa del inversionista cuando se le informó que, para aprobarle su crédito, en vez de solicitar el equivalente de 80,000 dólares para financiera su capital de trabajo, debía hacerlo por el equivalente a 140,000 dólares. Sólo así se le podría aprobar el crédito.

Lo anterior se justificó por el hecho de que el banco compartió mi conclusión de que, al prestarle sólo 80,000 dólares, su cliente iba a enfrentar serios problemas de operación a mediano plazo, por cuando sus requerimientos reales eran mayores que los incorrectamente calculados.

De esa forma, el banco evitó que un proyecto rentable fracasara por falta de recursos para la operación, logrando, al mismo tiempo, velar por sus propios intereses en resguardo de los recursos otorgados en crédito.

La lista, sin embargo, no termina aquí. Las incoherencias se manifiestan en todos los ámbitos de la formulación. Por ello, uno de los aspectos centrales que se intentará cubrir en los siguientes capítulos, de este libro será justamente el relativo a cómo estructurar un cuadro coherente de datos para incorporarlo en la formulación del proyecto.

EVALUACIÓN DE OPCIONES EN LA FORMULACIÓN

Todo lo anterior guarda relación con la necesidad de mejorar la calidad de los estudios de proyectos, que si bien ha manifestado un notable desarrollo en los últimos tiempos, son aún susceptibles de perfeccionarse, sobre la base de la permanente búsqueda de optimizar el trabajo profesional que nos motiva.

Pero la optimización del proyecto no debe concentrarse solamente en el trabajo del evaluador, ya que una de las principales responsabilidades que a éste le competen -- y que no siempre se asume correctamente—es la optimización del uso de todos los recursos para emplear en el proyecto.

Así, entonces, al preparar un proyecto será preciso efectuar una serie de estimaciones de lo que se espera sean a futuro los beneficios y costos que se asocien a su operación, lo que requerirá que previamente se tome un sinnúmero de decisiones respecto a casi la totalidad de las características que debiera tener el proyecto, las cuales influirán directamente sobre el resultado de la evaluación.

Decisiones típicas a este respecto son, por ejemplo, las relativas a si la empresa que crearía el proyecto deberá elaborar o comprar sus insumos, si deberá arrendar o comprar las obras físicas en que se localizará su operación, si deberá usar una tecnología intensiva en capital u otra intensiva en mano de obra, si deberá hacer el transporte en medios propios o en ajenos, si deberá funcionar en una o en más plantas, si deberá implementar sistemas manuales o mecanizados de información, si deberá contratar segundos turnos u optar por remunerar sobre tiempos, etcétera.

NUESTRO COMETIDO

En las páginas siguientes se intentará contribuir al perfeccionamiento del trabajo que quienes deban medir la rentabilidad de una inversión, llamando la atención no sólo sobre los errores más frecuentes en la confección de los proyectos, tanto por omisión como por incorrección, sino que también sobre aquellos aspectos que conllevan una optimización de los recursos invertidos al incluir correctamente una mayor cantidad de variables en la estimación.

Es probable que el lector concluya, como resultado de la revisión de este libro, que el evaluador tendrá que enfrentar su tarea con mucho mayor trabajo del que estimaba antes de iniciar su lectura. Es más, esperamos que sea así. Sin embargo, la importante responsabilidad que asume en esta tarea asignada de recursos ante los distintos agentes

involucrados —el inversionista que enfrenta esta empresa arriesgando su capital, tiempo y prestigio; la institución financiera que respaldará al proyecto con parte de los recursos financieros requeridos, y la propia sociedad en general, que se beneficiará con los bienes u servicios que provea una correcta asignación de esos recursos—justifica sobradamente ese mayor esfuerzo.

Tal mayor esfuerzo deberá aplicarse, como fin último y prioritario, en la búsqueda de la mejoría del bienestar de las personas. Ello, en conformidad con los grandes desafíos que impone el desarrollo, sin perder de vista que es la técnica la que ha de estar al servicio del hombre, y no viceversa.

CAPITULO II

Limitaciones al trabajo del evaluador: las variables del entorno

UN IMPOSIBLE: LA PREDICCIÓN PERFECTA

Dentro del área de la administración, uno de los desafíos profesionales más difíciles es aquél que se relaciona con la medición de la rentabilidad de un proyecto.

Lo anterior es así, básicamente, por que se exige que con gran anticipación el evaluador determine la conveniencia de implementar una cierta iniciativa de inversión, mediante la estimación de los costos y beneficios que se ocasionan a la puesta en marcha y futura operación de un proyecto, incluso son que de él dependa ni siquiera la fecha en que se iniciará su construcción.

Hace unos años, por ejemplo, evaluamos la conveniencia económica de establecer y operar un puerto marítimo de embarque de fruta para la exportación, el que debía entrar en operación en diciembre de 1989. el proyecto se justificaba ampliamente, por situarse en una zona relativamente cercana a las plantaciones, perdiendo un considerable ahorro en el costo de los fletes, por la disminución en varios cientos de kilómetros de la ruta por la que debía transportarse la fruta al puerto de embarque de alternativa. A fines de 1990, el proyecto no sólo no había iniciado su etapa de construcción, sino que, además, la construcción de un puerto de otro inversionista se encontraba en sus primeras etapas en un punto cercano de la misma localidad.

Probablemente, si se llevase a cabo una actualización del estudio de la rentabilidad del mismo proyecto, se revelaría la inconveniencia de su implementación, a cauda del nuevo competidor no previsto en la evaluación original.

LA FINALIDAD DEL ESTUDIO DE UN PROYECTO

En este proceso de anticipación, el evaluador y su cliente deben tener muy en claro qué se puede esperar de su trabajo, y no caer en la tentación de sobredimensionar las capacidades de la técnica comúnmente denominada de preparación y evaluación de proyectos.

No son pocos los casos de quienes hoy lamentan haber tomado una decisión, ya sea invertir sus recursos o de prestarlos, basados principalmente en el resultado de una evaluación de proyectos.

Para nosotros, el resultado de un estudio de viabilidad no es sino un antecedente más para ayudar al inversionista a tomar la decisión de llevar a cabo una determinada iniciativa de inversión. Su carácter anticipado determina, incluso, que su resultado pueda estar totalmente equivocado. Y no necesariamente por un mal trabajo del evaluador, ya que el resultado dependerá del comportamiento de las variables que lo condicionan y sobre las cuales éste no tiene ningún grado de control.

LA INESTABILIDAD DE LA NATURALEZA

¿Cree el lector que se puede afirmar que el proyecto de poner los gaviones de protección en la rivera del río Mapocho, en la capital de Chile, para evitar los enormes costos que ocasionaban sus desbordes fue mal evaluado? Como se recordará, hace algunos años las lluvias, que excedieron todo promedio histórico, provocaron que este río se saliera de su cauce, ocasionando enormes daños a la propiedad. El argumento más escuchado en esa oportunidad era que tales fenómenos de la naturaleza ocurrían cada cuarenta años. Sin embargo, el año siguiente la situación se repitió, con nuevos grandes daños a los bienes públicos y privados.

Un año después —y empleando recursos que podrían haberse destinado a solucionar otros problemas urgentes de la zona afectada— se inauguraron los gaviones que garantizaban la protección contra esos desbordes; sin embargo, los tres años siguientes se caracterizaron por ser casi los más secos de la historia. El evaluador del proyecto que recomendó dicha solución, ¿habría podido ver los cambios que sucederían en la naturaleza? Obviamente no.

UN PROBLEMA DE LA EVALUACIÓN DE LOS PROYECTOS DE PREVENCIÓN DE RIESGO

En relación con este caso, es conveniente señalar la confusión que se crea a veces en los proyectos de prevención de riesgos: si las lluvias hubieran intentado desbordar al río Mapocho y los gaviones lo hubieran impedido, lo más probable es que ese evaluador recibiera sólo elogios, aunque esas mismas lluvias hubiesen causado daños a la propiedad en otros lugares. Similar es el caso de aquellas personas que se lamentan por haber suscrito y apagado un seguro para su automóvil cuando aquel vence sin haber sido ocupado, pero que se alegran de su inversión cuando ocurre un siniestro.

Cualquiera sea el caso, la inversión, si ya se efectuó, tiene carácter de irrelevante para la decisión, y por lo tanto no debiera esperarse la ocurrencia del siniestro para poder justificar el proyecto. La importancia de este aspecto justifica que más adelante se dedique un capítulo completo a este tema.

LA REPERCUSIÓN DE LOS CAMBIOS DE LAS POLÍTICAS ECONÓMICAS

Pero no es sólo la naturaleza la variable del incontrolable que puede llevar al fracaso al mejor de los evaluadores. Hace unos años, se nos pidió evaluar la conveniencia de instalar

y operar una fábrica de envases de aluminio. La principal competencia que enfrentaba el proyecto radicaba en el envase de hojalata, cuya materia prima era origen nacional. El aluminio que requería el proyecto al que nos enfrentábamos era en su totalidad importado.

Si bien los resultados de la estimación inducían pensar en la conveniencia de recomendar su pronta implantación, había un factor que nos hizo dudar acerca de ello: la posibilidad de que un elemento incontrolable por cualquier evaluador —como era en ese momento el tipo de cambio—fuera modificado por una decisión de la autoridad económica de turno, lo que a su vez podía originar que tanto el costo como el precio de venta del envase de aluminio se incrementasen a corto y mediano plazo hasta niveles no competitivos respecto al envase de hojalata.

Probablemente, varios evaluadores habrían intentado la aplicación de alguno de los diversos modelos que provee la teoría para proyectar el comportamiento futuro del tipo de cambio para incorporarlo en la proyección de los costos y beneficios. Sin embargo, concientes de que el nivel que esta variable alcance depende de una decisión política (de política económica) externa al proyecto, se decidió reconocer nuestra incapacidad para proyectar técnicamente su valor esperado y, en su reemplazo, ofrecer un análisis de sensibilidad a dicho evento.

El tipo de cambio vigente en ese momento se mantenía fijo desde hacía bastante tiempo y, aunque la autoridad económica señalaba por todos los medios que no se modificaría, muchos dudaban que pudiera sostenerse constante por más tiempo. Al sensibilizar el resultado del proyecto, se determinó, que éste resistiría una modificación en el tipo de cambio que no sobrepasara un incremento de 25 por 100 y muy pronto sobrepasará el nivel máximo que podía resistir su evaluación.

Si el proyecto se hubiese implementado, ¿podría alguien calificar de ineficaz nuestro trabajo? Creemos que no, por cuanto en el informe se había reconocido que el resultado dependía de una variable que no creímos posible de proyectar con los modelos que provee la teoría económica.

Por otra parte, la permanencia en su cargo de la autoridad económica que dicta las políticas depende, además de la desconfianza del gobernante de turno, de la propia permanencia de éste en el poder. ¿Pueden los evaluadores proyectar la tendencia política, primero, y económica después para un país durante todo el período de evaluación del proyecto, que generalmente es de diez años?

LOS ALBURES DE LAS RELACIONES INTERNACIONALES

En el caso de los envases, no se había alcanzado a implementar el proyecto. Pero hubo otro caso en el cual un proyecto recién puesto en marcha con grandes expectativas se debió abandonar por un cambio en otra variable incontrolable, la planta elaboradora de papel para periódicos que utilizaría la cáscara o bagazo de caña de azúcar como materia prima. Las bondades del proyecto residían en el hecho de que el bagazo se botaba o incluso quemada para deshacerse de él. Por eso, el proyecto se evaluó con un costo de la materia prima correspondiente sólo a su transporte.

Se cuenta al respecto que los productores de caña de azúcar, que debían pagar para quemar o botar su bagazo, empezaron a notar que cada vez más camiones esperaban para cargar y llevarse esta aparente basura. Esto los motivó a poner un precio inicialmente bajo a este residuo, para luego ir subiéndolo en la medida en que había compradores dispuesto a

pagar esos precios, hasta llegar a un nivel de equilibrio tal, que la fábrica de papel de periódico pasó a ser un negocio no rentable. La causa de estos aumentos previstos se originó en un primer aumento del precio del petróleo, que obligó a buscar sustitutos más baratos para generar cierto tipo de energía. El bagazo de la caña de azúcar resultó ser de alto poder de combustión.

La misma crisis del golfo pérsico en 1990 hizo subir los precios del petróleo a niveles que ningún evaluador proyectó en su trabajo. ¿Qué habrá pasado en ese proyecto de tan alta rentabilidad del proyecto en su trabajo. ¿Qué habrá pasado en ese proyecto de tan alta rentabilidad que, para evitar el impacto de las sequías sobre la generación de energía por vía hidroeléctrica, propuso implementar plantas termoeléctricas que consumían sobre mil litros de petróleo por minuto?

EL AVASALLANTE CAMBIO TECNOLÓGICO

También es clarificador en este sentido el caso en que se inició la construcción de una planta para procesar pectina a partir de la cáscara de limón que desechaban las fábricas de aceites esenciales de limón. La alta rentabilidad del proyecto, sustentada principalmente en el hecho de que sólo un país latinoamericano producía pectina, aunque en cantidades inferiores a su consumo, por lo que estaba también obligado a importar la diferencia, así como por la gratuidad de la materia prima, condujo a que recomendáramos su más pronta implementación. Sin embargo, antes de finalizada la construcción el proyecto debió ser reevaluado, ya que surgió un imprevisto determinante en su rentabilidad futura: la aparición de un sustituto sintético que reemplazaba a la pectina natural en gran cantidad de usos.

Los evaluadores originales no fuimos capaces (¿alguien lo hubiera sido?) de visualizar este nuevo producto sustituto. La cuantía de los recursos involucrados en investigaciones tecnológicas induce a las empresas a mantener una gran reserva tanto acerca del tema investigado como de sus estados de avance.

Considérese al respecto que el descubrimiento de una nueva molécula tiene un costo estimado promedio de 150 millones de dólares, pero copiarla sólo cuesta algunos miles de dólares. De aquí la importancia que ha adquirido la discusión del tema de las patentes farmacéuticas suscitada por reclamaciones estadounidenses.

Y ASÍ SUCESIVAMENTE

Todos los casos anteriores muestran aparentes fracasos motivados por cambios en las variables del entorno del proyecto; la naturaleza, la política económica los cambios tecnológicos o los precios relativos internacionales. Son muchos los casos que se podrían agregar para gratificar situaciones similares ocasionadas por cambios en la normativa legal o tributaria (un nuevo impuesto al patrimonio después de una catástrofe nacional); en el marco institucional (el cambio en los precios relativos de las propiedades por el traslado de la sede del Congreso a otra ciudad o el cambio de la capital de un país); cultural (la mayor demanda de lectura al disminuir la tasa de analfabetismo), etcétera.

Sin embargo, no todos los cambios incontrolables son contrarios a la rentabilidad de un proyecto. Nadie puede considerar positivo un terremoto, pero (planteándolo frívolamente) si destruye a las empresas competitivas aumenta la rentabilidad del proyecto. La

aparición de un sustituto más barato de la pectina es muy conveniente para los consumidores de ella. En todo cambio de las políticas económicas, alguien gana lo que otro pierde.

La dificultad de poder prever en qué áreas del entorno es posible esperar cambios es incrementada por la imposibilidad de saber si ellos serán beneficiosos o contrarios a los intereses del proyecto.

Por otra parte, el evaluador se enfrenta además al problema de proyectar los ingresos y egresos de un proyecto del cual no controla ni siquiera la fecha en que se llevará a cabo. Muchas veces, mientras el inversionista está en la etapa de la búsqueda de las fuentes de financiamiento para iniciar las inversiones proyectadas, los cambios en el entorno pueden dejar obsoleto el resultado del estudio.

FORMULACIÓN DE PROYECTOS: UN DESAFÍO INELUDIBLE

Después de leer estas líneas es probable que el lector quede con un resabio de desconfianza en las técnicas de evaluación de proyectos. No ha sido éste el objeto de este capítulo, sino, por el contrario, se persigue eliminar el exceso de confianza que hoy apreciamos en muchos analistas en relación con sus resultados.

Se ha mostrado cómo el entorno constituye una variable incontrolable que el evaluador no debe subestimar. En los capítulos que siguen esperamos ayudar con nuestros comentarios a abordar correctamente el estudio de propuestas de inversión.

Para estos efectos se analizarán los requerimientos mínimos que deben observarse en el estudio económico de un proyecto, la forma de incluir un análisis de su entorno incontrolable, los errores más frecuentes en su formulación y evaluación, y los criterios más eficaces para efectuar correctamente un estudio en diferentes áreas de la actividad económica nacional, entre otros aspectos.

CAPITULO III

¿Viabilidad o factibilidad?

Es frecuente encontrar en estudios de proyectos la denominación genérica de <<factibilidad técnica económica>>, para titular el informe final de cualquier evaluación de un proyecto de inversión.

Si bien en el lenguaje diario se usan distintamente los términos viabilidad y factibilidad, es conveniente aclarar la necesidad de su diferenciación, especialmente si se considera que el segundo corresponde a una etapa del ciclo del proyecto que tiene repercusiones importantes en los procedimientos y alcances que se exigirán al trabajo del evaluador.

LAS DIVERSAS VIABILIDADES

Para recomendar la aprobación de un proyecto de inversión, el evaluador se enfrenta con tres viabilidades principales que investigar, entendiendo por viabilidad la <<posibilidad de>> o la <<conveniencia de>> realizar un proyecto: la viabilidad técnica, la viabilidad legal y la viabilidad económica.

Hay una cuarta viabilidad que es determinante en la realización de un proyecto, pero que no se incluirá en este texto: la viabilidad política, que corresponde a la intencionalidad, de quien debe decidir, de querer o no implementar un proyecto independientemente de su rentabilidad. Por ejemplo, por razones estratégicas o humanitarias, o simplemente por convenir a determinados intereses que pudieran estar en juego, podrá aceptarse un proyecto que muestre una rentabilidad negativa o rechazarse otro que tenga una rentabilidad positiva.

Sin embargo, el resultado de varios estudios realizados a causa del hecho de que muchos proyectos que al ser evaluados mostraron una rentabilidad positiva y que una vez implementado terminaron en un fracaso, nos lleva señalar la necesidad de incorporar una nueva viabilidad: la gerencial.

La falta de capacidad de gestión explica, lamentablemente, el fracaso de muchas iniciativas de inversión que, en otras circunstancias, pudieron haber proporcionado importantes ganancias a los inversionistas que las llevaron a cabo.

Una de las pocas formas de que se dispone para estudiar la viabilidad gerencial es la calidad del proyecto que se elabora. Si bien esta etapa tan relevante el inversionista demuestra incapacidad para hacer un buen estudio o para hacerse asesorar por un adecuado equipo de evaluadores, podría fácilmente presumirse que, una vez implementado su proyecto, mantendrá esta incapacidad.

LA VIABILIDAD TÉCNICA

La viabilidad técnica, que siempre debe establecerse con la ayuda de los técnicos especializados en la materia, busca determinar si es posible física o materialmente <<hacer>> un proyecto. Tal tarea corresponde a dichos especialistas y no puede ser asumida con responsabilidad por el evaluador económico del proyecto. Por ejemplo, sólo los expertos pueden, en sus respectivas áreas de especialidad, determinar si materialmente es posible construir un puerto en determinado lugar, obtener pectina del limón, producir papel de diario usando el bagazo de la caña de azúcar o reconvertir el plástico de desecho para utilizarlo como materia prima de otros productos plásticos.

Considérese al respecto lo absurdo que sería el hecho de que el evaluador económico tuviera que medir, con traje de buzo, si la profundidad del mar permite recalcar a las naves que habrían de usar el futuro puerto; o que en laboratorio haga pruebas para precisar el contenido de pectinas de cierto tipo de limón; o que mida la temperatura de las aguas para determinar la posibilidad de instalar un proyecto de piscicultura.

LA VIABILIDAD LEGAL

La viabilidad legal, por otra parte, se refiere a la necesidad de determinar la inexistencia de trabas legales para la instalación y operación normal del proyecto.

Al igual que en la viabilidad técnica, su realización corresponde a los expertos respectivos, y en ella el responsable de estudiar su conveniencia económica tiene poco o nada que decir.

Sin embargo, a diferencia del caso de la viabilidad anterior, el evaluador puede, mediante una investigación exploratoria, definir el marco de restricciones legales que enfrentará al proyecto, muchas restricciones se encuentran hoy día, por ejemplo, en los planos reguladores para construir determinados tipos de edificios para el cultivo marino en zonas que pudieran considerarse estratégicas para la defensa del país.

LA VIABILIDAD ECONÓMICA

En el estudio de la viabilidad económica se pretende definir, mediante a comparación de posbeneficios y costos estimados de un proyecto, si es recomendable su implementación y posterior operación.

En la concreción de esta viabilidad se reconocen tres etapas o niveles en que se clasifican los estudios de acuerdo con su profundidad y con la calidad y cantidad de información utilizada, siendo la última de tales etapas la de factibilidad. En este punto centraremos este capítulo, en nuestro intento de clarificar los conceptos, técnicas y metodologías acerca de la formulación y evaluación de proyectos.

COMO CONCEBIR UN PROYECTO

El estudio de un proyecto surge como respuesta a una <<idea>> que busca ya sea la solución de un problema o la forma para aprovechar una oportunidad. Un análisis preliminar de la situación debe permitir un juicio, también preliminar, de la posibilidad de concretar la idea en una acción.

Reconociendo la responsabilidad del evaluador de proyectos en la búsqueda de la optimización del uso de los recursos de los inversionistas, una tarea fundamental en esta etapa es la identificación de posibles opciones de solución.

La evaluación de cada una de ellas posibilitará elegir la más conveniente para el inversionista o la sociedad. Obviamente, una alternativa importante podrá ser siempre no hacer nada. Esto se conoce como la situación base o situación sin proyecto.

Es usual, entre inversionistas e incluso evaluadores de proyectos, es estudiar una iniciativa de inversión definida previamente, sin considerar la posibilidad de que exista una mejor solución.

LA IMPORTANCIA DE IDENTIFICAR OPCIONES

En una ocasión, por ejemplo, se presentó a un organismo financiero internacional un proyecto que demostraba la conveniencia económica –tanto privada como social—de construir y operar un nuevo servicio de salud en una localidad rural donde el hospital existente no tenía capacidad para atender la demanda de la zona. Al revisar los antecedentes proporcionados, se pudo detectar que la capacidad de este establecimiento debía ser suficiente para cubrir las demandas de servicios de salud de la población de la zona, de acuerdo con cualquier análisis y comparación de estándares.

Esta situación hizo necesario llevar a cabo un estudio que posibilitó determinar que el grado de conocimiento que los médicos tenían de sus pacientes, casi al extremo de amistad, los inducía a que, por consideraciones distintas de las exclusivamente médicas, los dejaran en recuperación por un período superior a lo usual en el resto del país.

En algunos casos, por tratarse de una anciana que al ser dada de alto volvería a sus esforzadas labores o, en otros, por tratarse de un niño al que harían trabajar en labores agrícolas antes y después de cumplir con su asistencia escolar. Sin siquiera cuestionar la posición de los médicos, de por sí valiosa por su carácter humanitario, pudimos demostrar, por la vía del absurdo, que a la comunidad le saldría más económico enviar esos pacientes al mejor hotel de la capital, con todos los gastos pagados, que construir un nuevo establecimiento hospitalario para ello.

Obviamente, había otras soluciones más eficientes, como, por ejemplo, ampliar sólo el pensionado, comprar una ambulancia para trasladar a los pacientes a otro centro asistencial ubicado a 220 kilómetros de distancia, o una solución máxima entre ambas posibilidades. La nueva evaluación mostró que la última era la más conveniente, tanto des de el punto de vista privado como social.

MEDIR LA RENTABILIDAD E LA IDEA

Para algunos autores, cada una de las posibilidades que se identifiquen pasa a constituir un proyecto y, como tal, deberá poder recomendarse el abandono, la implementación inmediata, la postergación o la profundización del estudio. Esta última se efectúa mediante una o más de las siguientes etapas, conocidas como preinversión: perfil, prefactibilidad o factibilidad. Y cada una de éstas se diferencia de las otras fundamentalmente en la cantidad y calidad de la información involucrada.

EL PERFIL DEL PROYECTO

La etapa de perfil es la más preliminar de todas. Su análisis es generalmente estático (compara, por ejemplo, los costos con los beneficios de un período tipo) y se basa en información secundaria, generalmente de tipo cualitativo, en opiniones de expertos o en cifras estimativas. Su objetivo fundamental es determinar si existen antecedentes que justifiquen abandonar el proyecto sin efectuar gastos futuros en estudios que buscan mayor y mejor profundidad.

Por ejemplo, en la evaluación de un proyecto para la construcción de un edificio destinado a locales comerciales, cuando el único existente mostraba, hasta ese momento, un evidente éxito, lo primero que se hizo fue averiguar cuántos permisos de construcción de edificaciones similares había otorgado la municipalidad respectiva. La magnitud de permisos aprobados o en etapa de estudio llevó a recomendar el abandono del proyecto, ya que, sin haber calculado siquiera los costos de construcción y los precios de venta, se pudo determinar que la competencia que enfrentaría al estar terminando lo hacía tremendamente arriesgado.

EL ESTUDIO EN NIVELES DE PREFACTIBILIDAD Y FACTIBILIDAD

Las etapas de prefactibilidad y factibilidad son esencialmente dinámicas; es decir, proyectan los costos y beneficios del proyecto a lo largo del tiempo y los expresan mediante un flujo de caja estructurado en función de criterios convencionales comúnmente establecidos.

La etapa de prefactibilidad proyecta los costos y beneficios sobre la base de criterios cuantitativos, pero sirviéndose de información secundaria (elaborada por terceros, como,

por ejemplo, tasa de crecimiento de la población revelada por el Instituto Nacional de Estadísticas, registros de importación del Banco Central, etcétera). La etapa de factibilidad, por el contrario, busca determinar la información de la fuente que la genera.

Una forma simple para explicar la diferencia entre ambos niveles de estudio dice relación con el caso de la compra de terreno para construir las instalaciones de la fábrica. Si la superficie del terreno requerido fuese de 3.000 metros cuadrados y el precio promedio del metro es de 25 dólares, en nivel de prefactibilidad se usaría como cifra de inversión al resultado del producto de ambos valores, es decir, 75.000 dólares. Sin embargo, en nivel de factibilidad deberá verificarse la existencia de un terreno con las dimensiones deseadas, ya que lo más probable es que no se obtenga uno del tamaño exacto necesitado. De esta forma, si el terreno más cercano al tamaño requerido tuviera 3.180 metros cuadrados, en nivel de factibilidad deberá considerarse una inversión de 79.500 dólares.

EL CASO DEL AEROPUERTO

Un caso que puede aclarar más esta diferenciación es el que le tocó realizar a la sociedad de profesionales ACEC Ltda., a petición de la Intendencia de la Tercera Región de Chile, para determinar la viabilidad de reabrir el aeropuerto Chamonate de Copiapó, allá por 1979.

El aeropuerto había sido cerrado en 1974, por tener una demanda inferior al mínimo que justificaba económicamente su operación. Para definir si las nuevas condiciones harían recomendable su reapertura, se procedió a proyectar la demanda mediante métodos estadísticos que tomaban el comportamiento histórico y lo proyectaban en función de alguna de las variables que se aceptaban internacionalmente como estándares para estos casos, como era la tasa de crecimiento de la población, entre otras.

Los antecedentes así procesados indicaban la existencia de una demanda superior a la mínima exigida para reabrir el aeropuerto. Sin embargo, las dudas que se mantenían sobre el procedimiento nos hicieron recomendar la profundización del estudio en el nivel de factibilidad.

En esta etapa se recurrió a la fuente misma de la información; es decir, mediante un estudio de investigación de mercado se procedió a investigar a los pasajeros de buses y vehículos que se movilizaban entre Santiago y Copiapó, quienes constituían la demanda potencial del proyecto.

El resultado indicó que no había la intención de cambio suficiente para asegurar la demanda mínima, ya que desde 1974 se había producido un cambio tan grande en la calidad, comodidad y horarios de los buses, que hacía menos atractivo el uso de otro medio opcional, como el avión.

UN CASO INMOBILIARIO

Otro caso a este respecto fue el de un proyecto de construcción y venta de un edificio de locales comerciales, oficinas y apartamentos de uso habitacional.

El nivel de prefactibilidad, el proyecto se evaluó basándose en informaciones de promedios de mercado (costo por metro cuadrado de las terminaciones de una obra gruesa construida, precio promedio de venta por metro cuadrado de locales de similar categoría en el centro de la ciudad, etcétera), obteniéndose un resultado que indujo a recomendar la implementación del proyecto.

Sin embargo, el cliente solicitó el estudio de factibilidad para tener más seguridad acerca de una inversión que recomendábamos efectuar. Para ello se contrató a una empresa experta en estudios de mercados habitacionales, a una de arquitectos y a otra de ingeniería, las que confirmaron, con un costo de varios millones de pesos, la recomendación preliminar.

Parece ahora fácil deducir que la etapa de factibilidad pudo haberse obviado. Sin embargo, durante la realización del segundo estudio, nadie podía estar seguro de cuán cerca o lejos estaría el resultado de los originalmente obtenidos.

EL PROYECTO EVALUADOR: FUENTE DE INFORMACIÓN

El problema se reduce, en consecuencia, al valor que cada una de las partes le da a la información. Mientras más aversión al riesgo tenga un inversionista, probablemente esté dispuesto a destinar más recursos a la realización de estudios más profundos.

La evaluación del proyecto, en este sentido, sólo ofrece más información de la que se tenía antes de hacer el estudio. La cantidad y calidad de ella dependerá del nivel de estudio que se contrate y de los recursos que estemos dispuestos a gastar en ella.

DE BONO, Edward; “Introducción” en: Más allá de la Competencia. México ED. Paidós pp. 11-24.

Introducción

En este libro deseo mostrar que la noción de competencia es una trampa peligrosa y seductora que limita y restringe el pensamiento empresarial. Todas las personas que dirigen una empresa necesitan ir más allá de la competencia hacia la *sur/petición*.*

La competencia es un concepto en boga vigorosamente impulsado por expertos tales como Michael Porter, de la Escuela de Negocios de Harvard. Pero cualquier escuela de negocios debe estar rezagada unos diez años en su pensamiento para ser creíble. Esto es así porque sus ideas deben ser inmediatamente aceptables, y las ideas del futuro no lo son. En consecuencia, tales conceptos no mejoran necesariamente la reputación de una escuela.

Todos conocemos el mercado y sabemos que para sobrevivir hemos de ser competitivos. Debemos poder competir con los japoneses, los alemanes y los taiwaneses. Si no podemos hacerlo, no sobreviviremos. Por lo tanto, ¿qué hay de malo en la competencia?

La paradoja está en que no se podrá ser verdaderamente competitivo si se intenta ser competitivo.

La palabra clave es <<sobrevivir>>. Sin duda usted deberá ser competitivo para sobrevivir. Los grandes minoristas como Sears deben reducir considerablemente sus costos para poder sobrevivir frente a otros minoristas tales como Wal-Mart con sus avanzados sistemas informativos y elevado número de ventas por metro cuadrado. Si sus costos y valores no están en línea, puede dejar de sobrevivir.

Pero cualquier organización que planifica tan sólo sobrevivir se hallará a la larga fuera del negocio. Sólo las organizaciones que planifiquen para el éxito sobrevivirán, mientras que aquellas que planifiquen para sobrevivir, fracasarán.

* El autor introduce aquí un neologismo (formado a partir de competición, es decir <<competencia>>), cuyo sentido explica en las páginas siguientes. Hemos considerado que conservar el neologismo era preferible a sustituirlo por una traducción literal (<<sobre/petencia>>, por ejemplo) escasamente eufónica. [T.]

Por lo tanto, la competencia es importante como parte de la <<línea de base>> para la supervivencia.

Imagine un jardín exótico en las afueras de Manila, en Filipinas. Una perfumada tarde de verano es el marco idóneo para un banquete de los Chevaliers de Tastevins. Se colocan antorchas encendidas entre los arbustos. Mozos vestidos de etiqueta llevan el primer plato de mariscos sobre sus cabezas. Le sigue una sopa clara en grandes boles de cerámica. Todos –me incluyo--, comenzamos a tomar la sopa. Verdaderamente es muy clara. En realidad no es sopa. Es simple agua hervida en grandes aguamaniles para lavarse los dedos después del plato de mariscos. El *error* es comprensible.

El agua es necesaria para la sopa; pero la sopa debe ser algo más que agua.

Del mismo modo, hay muchos elementos necesarios, pero no suficientes, para que la empresa sobreviva (por ejemplo, el control de costos), tal como el agua para la sopa. La competencia es uno de los factores primordiales para que la empresa sobreviva; pero no es suficiente. Uno de los graves errores que comenten numerosos ejecutivos consiste en creer que la competencia es la clave para el éxito. No lo es. La competencia apenas forma parte de la línea de base de la supervivencia. El éxito exige ir más allá de la competencia hacia la *sur/petition*.

Sur / petition

Hay una grave sobreproducción de aproximadamente un 25% en la industria automotriz europea. En una ocasión me encontraba dando un seminario para el departamento británico de comercialización de Ford, en sus oficinas de la afueras de Detroit, y estábamos tratando la competencia en el mercado europeo.

Sugerí que Ford comprara una empresa denominada NCP (Nacional Car Parks), que era propietaria de la mayoría de los estacionamientos del centro de las ciudades en todo el Reino Unido. Si NCP se convertía en una empresa de Ford, podría colocarse un cartel en todos los aparcamientos que indicara que sólo podían utilizarlos automóviles Ford.

Un automóvil., sostuvo, ay no es sólo un montón de ingeniería. A un vecino que se jacte de tener un montón de ingeniería mejor que el suyo, bastará con señalarse que usted puede estacionar en la ciudad, y él no. Poder aparcar forma parte del <<valor integrado>> de un automóvil si de debe conducir en una ciudad. Lo mismo ocurre con la posibilidad de revenderlo, el hecho de contar con un buen servicio y el de tenerlo asegurado.

Me han dicho, por ejemplo, que es imposible asegurar un Mercedes o un BMW si se vive en el Bronx, en Nueva York. No importa lo buena que pueda ser la ingeniería; Si usted no puede asegurar un automóvil, probablemente no lo comprará.

Por su puesto, Ford no aceptó mi idea. La empresa afirmó que siendo una compañía industrial no era asunto suyo comprar aparcamientos. En el futuro, algunos empresarios probablemente compren o construyan estacionamientos y luego soliciten a los coreanos que fabriquen automóviles de marca privada para ellos. Venderán, estacionarán y revenderán automóviles. La fabricación será sólo un servicio para este centro de ganancias y se reducirían sus márgenes.

En Estados Unidos, las ventas de automóviles han estado decayendo desde hace algún tiempo. En cinco años (de 1986 a 1991) bajaron de 16,23 a 13,5 millones de unidades aproximadamente. Como todos sabemos, la clásica respuesta competitiva ha sido reducir drásticamente los precios y ofrecer descuentos por pago al contado. Sorprendentemente,

los competidores hacen lo mismo. Usted puede lograr el efecto deseado, pero los compradores se acostumbran a los descuentos y los aguardan antes de volver a comprar.

Mi sugerencia a Ford de comprar los estacionamientos del centro de las ciudades es un ejemplo de lo que denomino *sur/petition*.

La palabra <<com-petencia>> proviene del latín y significa <<buscar juntos>>. Significa <<optar por correr la misma carrera>>. Podríamos escribirla <<(con)petencia>> para mostrar que todos los competidores están en la misma carrera. La palabra *sur/petition* significa <<buscar sobre>>.en lugar de optar por correr la misma carrera, los competidores eligen su propia carrera. La barra en la nueva palabra <<*sur/petition*>> indica la noción de buscar <<sobre>>, tal como 2/3 indica dos sobre tres.

Además, la *sur/petition* se refiere a la creación de >>monopolios de valor>>. Algunas maneras de lograr un monopolio son ilícitas, pero los monopolios de valor no lo son. Para sobrevivir se necesita competencia; para tener éxito se necesita *sur petition* y creación de monopolios de valor.

La diferencia entre competencia y *sur/petition* se ilustra en la figura I.1 En lugar de correr la misma carrera, usted crea la suya. En lugar de buscar <<junto con>>, usted se lanza a buscar <<sobre>>.

Valores integrados

¿Cómo se crean monopolios de valor? Esto es precisamente lo que discutiré a lo largo de este libro. Los monopolios de valor están impulsados por conceptos están a su vez impulsados por una creatividad seria.

Para comprender los monopolios de valor, deberá advertir que ha habido tres fases en la empresa.

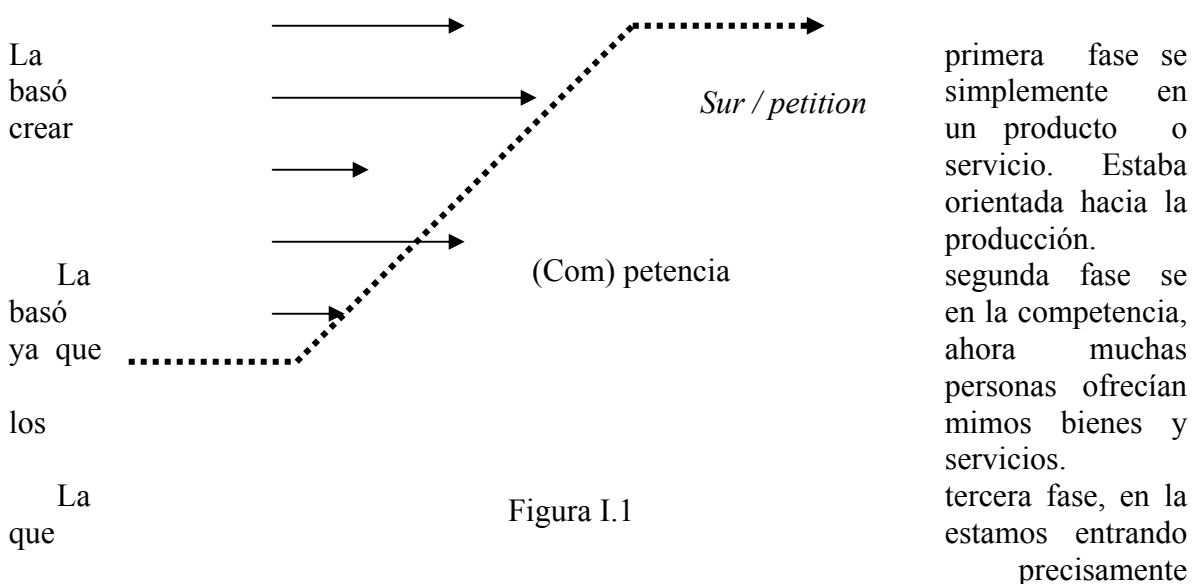


Figura I.1

ahora, se basa en los valores integrados. Ya no vivimos en un mundo de valores simples en el cual un automóvil es una pieza de ingeniería.

Un buen ejemplo de valores integrados lo da Ron Bárbaro, en una época presidente de la Prudential Insurance Company en Canadá. Utilizando algunas de mis técnicas de pensamiento lateral (creatividad seria), dio con la idea de <<beneficios en vida>>. Este es un cambio significativo en el muy tradicional negocio del seguro de vida, que ha permanecido invariable durante 120 años. En el seguro de vida habitual el importe asegurado se paga a la familia o a otros beneficiarios después de la muerte del titular. La idea de Ron Bárbaro fue la de pagar el 75% de ese importe al titular en caso de que se le diagnosticara una enfermedad potencialmente fatal (tal como cáncer o sida). Esto significaba que el dinero estaba disponible para cuidados adicionales o atención médica. La idea tuvo una excelente acogida y fue una de las razones por las que Ron –Bárbaro fue rápidamente ascendido a presidente Prudential en los Estados Unidos.

Este es un excelente ejemplo de valores integrados porque integra el seguro de vida a las vidas y valores de las personas. Hay gente soltera, otros que se divorcian y se separan, niños que crecen y se vuelven autosuficientes, etc. para algunos, muchos de los objetivos originales del seguro de vida dejan de tener vigencia. Al mismo tiempo, hay factores tales como el sida y los gastos de asistencia médica que crean nuevos objetivos.

Bárbaro produjo un simple cambio de conceptos. Tradicionalmente, se relaciona el seguro de vida con la muerte; si concepto enfatizó la vida y tuvo el coraje y el impulso para imponer esta idea.

Conceptos y creatividad

He trabajado en cuarenta y cinco países y he descubierto que la comunidad empresarial que está más insatisfecha con los conceptos en la de Estados Unidos. Hay una impaciencia con respecto a los conceptos y un deseo urgente de que se den herramientas <<directas>>. Los norteamericanos quieren hacer en lugar de pensar.

Esto no es algo nuevo. Durante mucho tiempo, Estados Unidos ha sido una sociedad pionera y, en una sociedad de este tipo, la acción recompensa mucho más que el pensamiento: es necesario limpiar algunas hectáreas más para pastoreo, perforar algunos pozos, abrir algunas tiendas más. La gran energía y falta de inhibiciones en Estados Unidos favorecen también la acción.

Sin embargo, hoy el mundo está más poblado. La acción impensada no será recompensada. Por otra parte, un cambio de concepto puede ser muy bien premiado. Un nuevo concepto es incuestionable la mejor y menos costosa manera de obtener un valor agregado de los recursos existenciales.

Sin embargo, hasta ahora, nuestro enfoque de desarrollo de concepto ha sido algo fortuito. Ha habido un enfoque <<yo también>>, según el cual esperamos que alguien lo desarrolle y luego nos introducimos con una copia de él. Además, hemos supuesto que, siendo personas inteligentes y capaces, daremos con los conceptos necesarios en el momento oportuno. Hoy esta actitud frente a los conceptos no es suficientemente buena.

En el futuro, tendremos que tomar los conceptos tan seriamente que montaremos departamentos especiales para <<Investigación y Desarrollo de Conceptos>> para la ingeniería intelectual que es necesario que se produzca. Se dice que General Motors gasta 5000 millones de dólares por año en investigación y desarrollo técnico y Du Pont gasta aproximadamente 1200 millones de dólares. Sin embargo, hoy los conceptos son aún más importantes que la tecnología y tendremos que tomar muy en serio su desarrollo.

Las escuelas de administración de empresas han enseñado siempre a sus estudiantes a analizar la información y a tomar decisiones. Hoy esto es totalmente inadecuado. El análisis de información y la toma de decisiones forman parte del aspecto de mantenimiento de la administración, forman parte del agua en la sopa. Ahora debe desplazarse el énfasis hacia el pensamiento conceptual, no como reemplazo del análisis de información y la toma de decisiones, sino sumándose a ellos. ¿Qué ideas se derivan del análisis de información? Por sí mismo, el análisis no puede dar nunca como resultado los conceptos que se hallan ocultos en la información. ¿Qué alternativas hay para la toma de decisiones? El análisis sólo puede ofrecer algunas de las alternativas; el resto debe producirse con un diseño creativo.

Desafortunadamente, el enfoque de la creatividad en Estados Unidos ha sido muy débil. Existe la errónea noción de que todos somos básicamente creativos y de que eliminando nuestras inhibiciones y temor al ridículo conseguiremos liberar esa creatividad innata. Este es el motivo de que los métodos creativos norteamericanos hayan sido desvariados y extraños.

Hoy sabemos, gracias al comportamiento de los sistemas auto otorgados de información (tales como la percepción humana), qué necesitamos hacer, conocemos que el cerebro no está diseñado para ser creativo y que para ello debemos utilizar algunos métodos que son >>no naturales>>. Hemos de comenzar a desarrollar métodos sistemáticos de creatividad seria. Lo discutiré más adelante, particularmente su relevancia para el desarrollo de conceptos.

A fin de concentrarnos en una creatividad seria, he montado un Foro Creativo Internacional para reunir a todas las personas de las principales empresas que están directamente relacionadas con el desarrollo y aplicación de esta creatividad. Entre los miembros fundadores se encuentran IBM, Prudential, Du Pont, Merck, Nestlé, British Airways (la más grande transportista internacional) y BAA (la más importante administradora de aeropuertos del mundo). Me alegra ver cómo las principales empresas están comenzando a tomar la creatividad de forma seria.

Valufactura

Los conceptos se refieren a valores. La *sur/petition* se refiere asimismo a valores. El éxito futuro de toda empresa se referirá a valores.

Siempre me ha extrañado que, aunque hablamos tanto sobre los valores, no tengamos una palabra específica para describir su creación y formación. Esa palabra nos permite una visión diferente y poder dedicar un esfuerzo mucho mayor a la creación específica de valores. La nueva palabra que propongo es <<valufactura>>, que defino como <<la creación y formación de valores>>. Existe una analogía con <<manufactura>>, que tiene que ver con la creación y producción de objetos.

La era de contracción

Mucha gente de negocios cree que hemos entrado en la era de contracción. Ha habido dos eras de expansión. La primera estuvo impulsada por la comercialización, la cual crearía necesidades y abriría mercados. La comercialización era la principal herramienta para quitarles la participación en el mercado a otros. La segunda era de expansión estuvo impulsada por el <<crecimiento devorador>>. Consistía en devorar a otras organizaciones a

fin de incrementar la propia participación en el mercado. Durante este período hubo toda clase de racionalizaciones sobre sinergias y tamaño de supervivencia crítica (generalmente brindados por bancos de inversión a los que les complacían las elevadas tarifas resultantes).

Las dos eras de expansión ya han pasado. El estado de ánimo actual favorece la consolidación. Deben dirigirse las adquisiciones. Los ejecutivos de las empresas están mirando hacia adentro. Hay reducción de costos y adelgazamiento. Hay despidos de personal. En cinco años CMS Energy Corporation redujo su fuerza de trabajo de 13.000 a 9500 personas. DEC (anteriormente Digital Equipment) desea reducir su nómina de 124.000 en el 6%. Hay desinversión de empresas que dan pérdidas en todas partes.

Como parte de esta mirada hacia adentro, hay un fenómeno en la reducción y el control de costos, y en la administración de calidad.

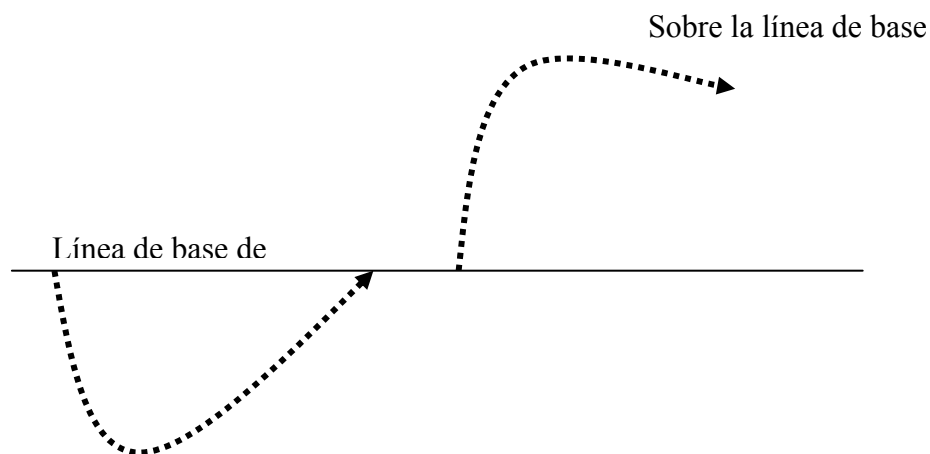
Todo esto es importante y necesario, pero no suficiente. Quizás usted se toma una aspirina cuando tiene jaqueca, pero no podrá sobrevivir con una dieta continua a base de aspirinas.

La figura I.2 nos muestra que si usted cae debajo de la línea de base de la competencia, necesitará regresar a ella. Pero estar sobre la línea de base no es suficiente. Debe ser mejor.

El propósito de la consolidación, contracción, calidad, etc. es brindar una línea de base firme para el emprendimiento. Finalmente deberá ofrecer los valores que los clientes desean. De nada sirve tener el producto de mejor calidad al precio más bajo si ese producto no ofrece in valor significativo.

Existe el serio peligro de que el actual énfasis de lo que yo denomino, <<manejo>> desvíe la atención de la esencia puede resumirse con las siguientes cuatro palabras que comienzan con <<C>>:

- Competencia: la calidad, eficiencia, eficacia y logro de lo que supuestamente ésta haciendo.
- Control: control de costos, estrategia y conocimiento de lo que sucede.
- Cuidado: cuidado del cliente, de la fuerza de trabajo, que es en última instancia su recursos más preciado, y de la comunidad (por ejemplo, las preocupaciones ambientales).
- Creatividad: el alma de la empresa. Sin creatividad, tiene un cuerpo sin alma. La creatividad brinda el valor que construye el propósito global de toda empresa.



Regresar a la línea de *Pag.- 327*

Formato de este libro

Todos recordaremos el poderoso efecto del libro *In Search of Excellence* (Petern y Waterman). La mayoría de la persona sabe que muchas de las empresas mencionadas en el libro ya no existen: por ejemplo, People Express. Sin embargo, el mensaje del libro persiste; las personas importan.

El método de estudio de caso para escribir un libro es engañoso. Quizás el autor tomo los casos para satisfacer sus necesidades. Las empresas que siguen la misa política que el sujeto del caso pero no tienen éxito son ignoradas. Generalmente es incorrecto atribuir el éxito de una empresa a un aspecto en particular de la política. El éxito depende de una multitud de factores interactuantes: condiciones del mercado, clima económico, competidores, personal, relaciones laborales, confianza del consumidor, etcétera.

En consecuencia, creo que el método de estudio de caso no es muy honesto. Más importante aún: no enseña más allá de algunos mensajes generales (por valiosos que éstos sean). Es más importante emitir el mensaje directamente y confiar en que la inteligencia del lector lo relacione con la circunstancia de su propia empresa.

Este libro ésta dividido en tres partes:

Parte I

- Una mirada a algunos de los hábitos fundamentales del pensamiento gerencial – eficiencia, resolución de problemas, mantenimiento y prevención de errores – y por que es necesario repensar aun esto s hábitos fundamentales. Si bien pueden haber sido valiosos en el pasado, hoy pueden ser inadecuados e incluso peligrosos.
- Una mirada a las recientes tendencias en el comportamiento gerencial: reducción de costos, desinversión, calidad, servicios al cliente, crecimiento devorados, etc. si bien algunas de estas tendencias n son excelentes, hay también aspectos negativos. Por ejemplo, la calidad no tiene suficientemente en cuenta calidad de qué y para quién.

Parte II

- *Sur/petition*: la diferencia entre competencia tradicional y *sur/petition*. Las tres fases de la empresa y la naturaleza de los valores integrados, y la totalidad del área de valores en general.

Parte III

- la sección práctica del <<como>>. Discusión sobre la valufactura, el diseño de conceptos y su investigación y desarrollo. La naturaleza de la creatividad seria.

Una nueva perspectiva

Una de las grandes ventajas que tengo como observador del pensamiento empresarial es que, si bien trabajo bastante en Estados Unidos y Canadá, también paso mucho tiempo en el resto del mundo: Europa, Japón, Corea y Sudamérica, inclusive en los países de la antigua URSS y China. A menudo es necesario mantenerse alejado del pensamiento empresarial norteamericano mantenerse alejado del pensamiento empresarial lugares ciertos hábitos son tan tradicionales que dentro de Estados Unidos ni siquiera se los considera como peculiares. Se necesita una perspectiva global.

También es esencial observar una amplia gama de empresas. Una persona que dirige con que lo haga –no podrá tener esta amplitud de acceso. He trabajado con empresas electrónicas, de informática, farmacéuticas, procesadoras de alimentos, de bebidas, químicas, líneas aéreas, bancos y muchas más.

Mi experiencia ha sido que el sector empresarial es el único que ha mostrado el mayor interés en el <<pensamiento>> como tal. El resto de los sectores, tales como el político y aun el mundo académico, está más preocupado por la discusión, por probar que un lado es correcto y el otro, incorrecto. El gobierno necesita pensar, pero sorprendentemente piensa poco. En la empresa existe un método de prueba de la realidad: los resultados contables. Usted podrá discutir que tiene razón hasta enrojecer, pero aun así puede caer en bancarrota.

La empresa maneja bastante bien el lado analítico del pensamiento. Pero se necesita mejorar el lado constructivo, creativo y conceptual. En el futuro, éste es al aspecto del pensamiento que será esencial para el éxito. Todo el resto es gobierno. Para estar seguro, deberá hacer que el uso sea satisfactorio y los conceptos brillantes no reemplazan el buen manejo. Pero esto no es suficiente. El agua es necesaria para la sopa, pero la sopa es algo más que agua.

Existe sólo una enfermedad grave en el pensamiento empresarial. Esta es la enfermedad de la arrogancia y la complacencia, sentir que <<somos buenos, que lo estamos haciendo bien y que lo sabemos todo>>. Afortunadamente para el resto del mundo, los japoneses están comenzando a padecer esta enfermedad.

Resumen

Los nuevos conceptos que se introducirían en este libro son los siguientes:

1. *sur/petition*, que va más allá de la competencia.
2. <<Valores integrados>> y la tercera fase de la empresa.
3. <<Valufactura>> y una nueva notación de valores.
4. Investigación y desarrollo de conceptos.
5. <<Creatividad seria>>.

EDWARD DE BONO

GISELLE O, Kniep Martín; “Portfolios del docente como responsable del desarrollo curricular y la evaluación” En: Portfolios del desempeño de maestros, profesores y directivos.

México Ed. Paidós pp 83-120

CAPITULO 5

Portfolios del docente como responsable del desarrollo curricular y la evaluación

El hecho de llevar un portfolio me ha permitido no ser tan exigente conmigo misma. Tengo tendencia a serlo porque soy

perfeccionista y siempre quiero hacer las cosas mejor. Las reflexiones del portfolio me han hecho ver los pasos positivos que he dado para mejorar mi labor docente. En ese sentido, me ha ayudado a pensar más y a practicar lo que predico. Como resultado de haber documentado mi propio progreso, no puedo negar las cosas que he hecho bien. Por lo tanto, estoy más conforme con mi modo de enseñar.

JEANNETTE ATKINSON, maestra de tercer grado
Escuela primaria Harris, Penfield, Nueva York

El rol más importante que cumplen los docentes en el aula posiblemente sea el de responsables del desarrollo de currículo y la evaluación diagnóstica. En este rol, los docentes diseñan experiencias de aprendizaje que posibilitan a los alumnos adquirir los conocimientos, las destrezas y las actitudes que necesitan tener. En muchos sentidos, este rol constituye un ejercicio de traducción, pues implica mucho más que la adquisición o transformación de información. Los alumnos llegan al aula con una amplia variedad de conocimientos previos sobre su mundo. Los docentes deben ayudarlos a adaptarse y asimilar nueva información de un modo que tenga sentido para cada uno de los alumnos.

El desarrollo curricular y la evaluación también entraña un cuidadoso proceso de selección. Yo no he conocido a ningún docente que no tenga demasiado que enseñar en un lapso de tiempo que nunca parece suficiente. Tal vez el mayor desafío que enfrentan los docentes sea discernir entre lo que es esencial y lo que no lo es.

Cuando los docentes toman conocimiento de la teoría del aprendizaje constructivista y de ciertas innovaciones como el lenguaje total, la evaluación alternativa y la integración curricular, tienen que desarrollar o adaptar su enfoque del currículo y la evaluación para reflejar sus nuevos conocimientos. La mayoría de los docentes no tiene tiempo para elaborar un nuevo currículo durante el año lectivo o el verano. Además, gran parte de su aprendizaje acerca de estas innovaciones tiene lugar mientras están enseñando. En algunos casos, los exámenes impuestos externamente parecen estar con conflicto con las nuevas prácticas, o los docentes reciben mensajes contradictorios de supervisores. De hecho, es muy notable que veamos grandes cambios en el modo en que los docentes enfocan el currículo y la evaluación, pero los cambios realmente se producen porque a los buenos docentes les interesa el desarrollo profesional.

Al estudiar los portafolios de desarrollo curricular y evaluación que han elaborado los docentes participantes en mis programas, he encontrado una interesante variedad de usos de las prácticas centradas en el alumno en materia de currículo y evaluación. En este capítulo, expondré algunos de ellos tal como podrían aparecer en un portafolio del docente como responsable del desarrollo curricular y diagnóstico.

Este tipo de portafolio describe las interrelaciones entre las prácticas del currículo, la enseñanza y la evaluación diagnóstica. Revela la capacidad del docente para aplicar conceptos y destrezas en la preparación de lecciones, actividades y evaluaciones que apuntan a determinados resultados claramente definidos. Un portafolio típico podría contener todos o algunos de los siguientes elementos:

- ☞ Una carta al lector en la que se describe el portafolio y sus objetivos.
- ☞ Una descripción del contexto en el que trabaja el docente

- ☞ Una lista de los resultados del aprendizaje de los alumnos previstos por el docente.
- ☞ Un diagrama o esbozo del currículo, que podría incluir los conceptos clave, las destrezas, los temas o las lecturas a abordar.
- ☞ Una descripción de las evaluaciones que empela el docente y su relación con los resultados deseados.
- ☞ Muestras de trabajos de los alumnos que ilustran las prácticas del docente en materia de currículo y evaluación.

En los párrafos siguientes ilustraré con mayor profundidad algunos de estos elementos. La figura 5.1 muestra el aspecto de un portfolio ejemplar. En el apéndice 6 se incluye la tabla de especificación completa –con los niveles rudimentario, incipiente, elaborado hasta ejemplar-- de un portfolio del docente como responsable del desarrollo curricular y la evaluación.

El contexto

Las prácticas relativas al currículo y la evaluación diagnóstica no pueden juzgarse independientemente de su contexto. Es decir, como lector, no puedo evaluar los méritos de una unidad curricular sino cuento con información concreta acerca del grupo de alumnos al que está dirigida, incluyendo la composición demográfica de la clase, las necesidades y capacidades de los alumnos, los conocimientos y destrezas previos requeridos para abordar la unidad y las condiciones bajo las cuales será impartida. En consecuencia, el portfolio debe incluir una descripción pormenorizada del contexto en el que trabaja el docente. El siguiente es un ejemplo de esta información contextual respecto de la escuela y el aula de quinto grado que enseña Cristianna Martindale:

Hace diecinueve años que ejerzo la docencia, y desde hace once años trabajo en la Escuela Primaria William Sydney Mount, del distrito escolar Three Village, en Stony Brook, Nueva York. La Escuela Mount es una institución ubicada en un barro residencial de clase media alta, a pocos minutos de distancia de la Universidad Stony Brook. Alrededor de 800 alumnos asisten a esta escuela.

El año pasado tuve a mi cargo una de las cuatro divisiones de quinto grado. Mi grupo era heterogéneo en materia de capacidad intelectual, pero todos los chicos que previamente habían sido evaluados como intelectualmente dotados asistían a una de las dos divisiones de quinto grado que hay en la escuela para estudiantes con dotes superiores. Los once varones y las diez niñas de mi clase provenían de hogares de similar nivel socioeconómico.

Cuatro de mis alumnos recibían clases de apoyo escolar. Tres de ellos asistían durante una hora, cinco veces por semana. Un alumno iba durante una hora, tres veces por semana. Este alumno también concurría a una clase de dicción de media hora, dos veces por semana. Además, otros dos alumnos recibían apoyo en el laboratorio de Matemáticas durante treinta minutos, dos veces por semana.

GISELLE MARTIN-KNIEP

FIGURA 5.1

Tabla de especificación de un portfolio ejemplar del docente como responsable del desarrollo curricular y la evaluación *

Generalidades: el portfolio cuenta la historia de un docente cuya meta es aunar las prácticas de la enseñanza y la evaluación, y orientarlas hacia resultados de aprendizaje significativos. El portfolio demuestra la excepcional capacidad del docente para aplicar conceptos y destrezas relacionados con el currículo y la evaluación en la preparación de clases, actividades y evaluaciones dirigidas a la obtención de determinados resultados claramente definidos. El docente ha incluido todos los esbozos de las evaluaciones, lo que le permite al lector ver los cambios y las mejoras efectuados para aplicar principios de diseño e incrementar la coherencia entre el currículo, la enseñanza y la evaluación.

Contexto: El portfolio proporciona un panorama claro y cabal del currículo del docente, incluyendo conceptos, temas, destrezas y evaluaciones. Las conexiones entre estos elementos son claras y evidentes. Los resultados en torno a los cuales se han diseñado en currículo son significativos, están precisamente expresados en función del aprendizaje de los alumnos y su descripción incluye concretos y observables.

Plan de evaluación (de dos a tres pruebas o un portafolio del alumno):

El plan se combina a la perfección con el ejercicio de la enseñanza por parte del docente. Las evaluaciones son oportunidades de aprendizaje y viceversa. Las descripciones de las evaluaciones son detalladas y suministran todo el contexto necesario, incluyendo el área de contenidos, abordada, los resultados previstos con sus correspondientes indicadores, los objetivos a alcanzar y una reseña detallada de las actividades previas y siguientes a las evaluaciones. El lector entiende cabalmente cómo se insertan las evaluaciones en el currículo y qué función cumplen en su utilización los docentes, alumnos, padres y otras personas. Por último, se especifica la cantidad de tiempo y esfuerzo que demandan.

Diseño del portfolio: El diseño del portafolio se orienta y pertenece al alumno, tiene un objetivo y un público reales y documenta el rendimiento, el esfuerzo y el progreso del alumno hacia la obtención de resultados significativos y claramente explicados. El lector ve con claridad cómo se inserta el portafolio en el currículo y qué función cumplen en su elaboración los docentes, alumnos y otras personas. El docente describe con precisión la forma de selección y reflexión de los alumnos respecto de sus trabajos. Para evaluar el portafolio, el docente aplica criterios explícitos de desempeño compartido, que los alumnos utilizan en el proceso de evaluación.

Descripción de las evaluaciones: Las evaluaciones son verdaderamente auténticas. Son tareas insertas en el currículo, substancialmente e integradoras, que requieren que los alumnos utilicen conocimientos previos, apliquen conocimientos y destrezas de una o más áreas de estrategias metocognitivas, hagan replanteos y correcciones. Las evaluaciones son valoradas por otros destinatarios fuera de la escuela, cumplen un propósito real y son lo bastante flexibles como para permitirles a los alumnos tener opciones y oportunidades de éxito. Son relevantes para la vida de los alumnos y toman en cuenta diferentes

necesidades y orígenes culturales.

las evaluaciones incluyen consignas que orientan las reflexiones de los alumnos sobre los productos y los procesos, y que pueden consistir en preguntas, listas de verificación o tablas de especificación. Las reflexiones e indicaciones demuestran que el docente comprende la necesidad de tomar en cuenta los diversos niveles de desarrollo y estilos de aprendizaje. el docente ha incluido trabajos de las evaluaciones con explicaciones claras acerca de cómo funcionan las evaluaciones con diferentes estudiantes. El docente ha señalado el grado en el que los alumnos en cuestión son representativos de la diversidad existente en el aula, permitiéndole al lector sacar conclusiones acerca del valor de la evaluación para una amplia gama de niveles de capacidad y conocimiento.

Estándares de desempeño: Los estándares de desempeño para las tareas correspondientes a la evaluación están claros para todos los interesados. El docente y los alumnos los determinaron conjuntamente y los articularon en forma de tablas de especificación, criterios de puntuación y/o ejemplares. Estos estándares sirven de guía a los alumnos para evaluar sus trabajos y fijarse metas para mejorar su desempeño. El diseño de la tabla de especificación se adecua perfectamente a la evaluación. Los alumnos pueden encontrar las destrezas correspondientes en cada nivel de la tabla de especificación y pueden usar los niveles descriptos para incrementar su aprendizaje y establecer metas concretas. Los niveles más bajos superan en cantidad a los más altos, lo que convierte la tabla de especificación en una excelente herramienta de andamiaje y de enseñanza. El nivel más alto está por encima del estándar buscado, pues hasta el alumno de mayor rendimiento es estimulado a mejorar. Además, la tabla de especificación incluye ejemplos reales para cada nivel, que los alumnos contribuyeron a identificar y evaluar.

* la tabla de especificación completa para el portafolio del docente como responsable del desarrollo y curricular y la evaluación aparecen en el apéndice 6.

A uno de ellos se le había detectado “dificultades de audición”, por lo que dos veces or semana se retiraba del aula media hora antes para ver a un especialista.

También tengo un alumno con un problema grave de salud. Su desempeño escolar era superior al promedio, pero faltó a clase tres periodos de tres semanas durante el año. Yo tenía que administrarle la medicación dos veces al día.

Además de la información precedente, este texto debe brindar un panorama claro del currículo del docente, que incluya conceptos, temas, destrezas y evaluaciones, así como las conexiones entre estos elementos.

RESULTADOS Y ESTÁNDARES DEL APRENDIZAJE

Los estándares del aprendizaje son las definiciones generales de los conocimientos, destrezas y actitudes que deben tener los alumnos al final de una unidad, de un año lectivo o incluso de toda su etapa escolar. En un portafolio del docente como responsable del desarrollo curricular y la evaluación, los resultados en torno a los cuales se ha diseñado el

currículo deben estar claramente expuestos en términos del aprendizaje de los alumnos, e incluir indicadores concretos y observables.

Los siguientes son algunos fragmentos del portafolio de Lisa Boernum que proporcionan el contexto en el que ha elaborado su currículo y sus evaluaciones diagnósticas. Lisa es docente de Educación Especial en el distrito escolar Sag Harbor, en Long Island Nueva York.

Como soy docente de Educación Especial, mis resultados se orientan al proceso, más que al contenido. Esta información me brinda la flexibilidad que necesito en una situación fluctuante. Los tres resultados que guían mi trabajo, junto con sus correspondientes indicadores, son los siguientes:

Resultado n° 1. Los alumnos emplearán una variedad e destrezas de comunicación para transmitir sus conocimientos, interpretaciones e ideas acerca de sus propias expectativas y niveles de desempeño.

Indicador n° 1. Describen con precisión sus puntos fuertes y débiles en los proyectos basados en el desempeño.

Indicador n° 2. Describen los conocimientos, los recursos o las destrezas necesarias para mejorar su trabajo (tareas en clase, tareas domiciliarias, apuntes proyectos).

Indicador n° 3. Emplean las destrezas de escuchar, leer, hablar y escribir para adquirir y transmitir una noción de sus propios niveles de desempeño y para fijar metas y estrategias destinadas a evaluar su nivel actual.

Resultado n° 2. Los alumnos usarán las evaluaciones grupales, las auto evaluaciones, sus conocimientos previos y las nuevas destrezas para elevar sus expectativas y niveles de desempeño.

Indicador n° 1. Emplean las estrategias de organización discutidas en clase (cronogramas, manejo del tiempo y manejo de los materiales) a efectos de identificar expectativas razonables para su trabajo y de fijar metas.

Indicador n° 2. Utilizan las tablas de especificación para los proyectos asignados y las anotaciones en la agenda diaria a efectos de fijar metas, supervisar el progreso y evaluar las expectativas personales.

Indicador n° 3. Al concluir un proyecto importante reflexionan sobre su progreso intelectual siguiendo las tablas de especificación, y las anotaciones en la agenda diaria.

Resultado n° 3. Los alumnos emplearán una variedad de recursos para adquirir conocimientos, destrezas y comprensión a efectos de evaluar con precisión sus niveles de desempeño.

Indicador n° 1. Emplean materiales de referencia (libros de texto, almanaque, atlas y obras literarias), conocimientos previos y actividades grupales con los pares para adquirir conocimientos, destrezas y comprensión de nuevos conceptos a efectos de llevar a cabo tareas de aprendizaje.

Indicador n° 2. Utilizan las tablas de especificación para evaluar su desempeño en cada tarea de aprendizaje.

Para contestar sus resultados e indicadores, Lisa incluye un calendario de los docentes de sexto grado, en el que estos enumeran los diferentes temas y asignaturas que enseñan cada

mes, junto con sus respectivos objetivos. El siguiente es un fragmento del calendario de sexto grado:

Enero

- 1) *estudio de Grecia*
 - a) *Construir templos.*
 - b) *Responder problemas.*
 - c) *Participar en el Concurso Panhelénico de preguntas y respuestas.*
 - d) *Participar en la asamblea*
 - e) *Intervenir en representaciones que escenifiquen mitos.*
 - f) *Escribir e ilustrar mitos infantiles.*
 - g) *Participar en un simposio.*
 - h) *Leer mitos griegos y completar el proyecto literario*
 - i) *Participar en el concurso de geografía de la revista Nacional Geographic.*
- 2) *Objetivos del estudio*
 - a) *Desarrollar destrezas de resolución de problemas*
 - b) *Analizar citas históricas.*
 - c) *Elegir la mejor decisión a través de evaluar los resultados.*
 - d) *Evaluar diversas formas de gobierno.*
 - e) *Relatar un mito griego en forma de obra de teatro.*
 - f) *Reconocer las necesidades del auditorio.*
 - g) *Ilustrar un texto escrito.*
 - h) *Puntuar diálogos.*
 - i) *Crear un conflicto en un relato.*
 - j) *Localizar países, regiones y accidentes geográficos*
 - k) *autoevaluarse.*

Este tipo de calendario les permite a los docentes determinar la coincidencia entre sus experiencias de aprendizaje y sus resultados y objetivos.

DIAGRAMA CURRICULAR

Un diagrama curricular es una matriz o un cuadro en el que se enumeran cronológicamente los contenidos clave, las destrezas, los recursos y hasta las evaluaciones utilizadas por el docente a lo largo del tiempo, los diagramas curriculares varían en cuanto a la cantidad de ítems incluidos, el nivel de especificidad en la descripción de esos ítems y el marco temporal considerado.

Andrea Gerstenblatt, maestra de segundo grado del distrito escolar de Farmingdale, en Nueva York, elaboró un diagrama curricular con referencias cruzadas entre los tópicos, los temas y los conceptos abordados, las destrezas y procesos que enseñó y las evaluaciones que utilizó. En la figura 5.2 aparece un fragmento de su diagrama.

El diagrama mismo es tan útil como la capacidad del docente para usarlo y aprender a partir de él. Andrea hizo ambas cosas, como queda demostrado en su siguiente comentario:

La parte más difícil en la elaboración de mi diagrama curricular fue el componente de la evaluación. Redacté muchas evaluaciones, pero no me convencían su validez ni su cantidad. Era evidente que estaba evaluando demasiado algunos tópicos mientras que otros no tenían ninguna evaluación. También me cuestioné si algunos de mis tópicos estaban suficientemente conectados con los resultados previstos. Algunos lo estaban y otros no. un problema con el que todos nos encontramos es que muchos de estos tópicos son requeridos en el currículo. Yo tengo que concentrarme en elaborar evaluaciones pertinentes a estos tópicos.

Uno de mis objetivos para el diagrama curricular era poner al día las evaluaciones y combinar algunas de ellas para crear otras más eficaces y en menor cantidad. A efectos de que resultarán más apropiadas, empecé por examinar los resultados previstos, específicamente en el área de la Lengua. Esto me ayudó a crear evaluaciones que abarcan distintas áreas de contenido y evaluaban más de una destreza o contenido, y a diseñar tareas convenientes para mis alumnos y para mí.

[...] El hecho de haber utilizado este año portafolios con fines de evaluación me han ayudado a perfeccionar mis evaluaciones. Los portafolios me dijeron más acerca de o que mis alumnos sabían y podían hacer que las “prolijas actividades al fin de cada unidad” que antes utilizaba.

Los diagramas les suministran a los docentes datos y pruebas relativos a lo que están tratando de lograr. También pueden servirles de apoyo para efectuar cambios en el Currículo, su enseñanza y su evaluación, basados en sus experiencias con los alumnos. Por esta razón, es mejor concebir los diagramas curriculares como “trabajos en marcha” que como documentos estáticos.

LAS EVALUACIONES DIAGNÓSTICAS

El portafolio del docente como responsable del desarrollo curricular y la evaluación también incluye un plan de evaluaciones que es compatible con el currículo y las prácticas de enseñanza de ese docente. Idealmente, las evaluaciones deberían ser tan congruentes con las oportunidades de aprendizaje que a un observador externo le resultaría difícil distinguir entre la evaluación y el currículo. En otras palabras, los alumnos deberían poder extraer información diagnóstica de las lecciones mismas. Las descripciones de las evaluaciones deben ser detalladas y suministrar todo el contexto necesario, incluyendo área de contenidos, resultados previstos con sus correspondientes indicadores, objetivos a alcanzar y una reseña detallada de las actividades previas y siguientes a las evaluaciones. Los lectores deben estar en condiciones de entender cómo se insertan las evaluaciones en el currículo y que función cumplen en su empleo los docentes, alumnos, padres y otras personas. Por último, debe especificarse la cantidad de tiempo y esfuerzo que estas demandan.

Figura 5.2

Fragmento de un diagrama de lengua de segundo grado.

Noviembre	Diciembre	Enero
-----------	-----------	-------

<p><i>Tópicos/conceptos</i> Comunidades: rural, urbana, suburbana. Poesía Fonética Símbolos cartográficos Estudio de un autor; Tomie de Paola. Léxico descriptivo</p>	<p><i>Tópicos/conceptos</i> Las ballenas. Navidad y <i>janukkah</i>* Poesía. Fonética</p>	<p><i>Tópicos/conceptos</i> Las ballenas Pueblos nativos de América del Norte Caligrafía Cuentos de hadas</p>
<p><i>Destrezas y procesos</i> Discusiones, Anotaciones en la agenda diaria. Lectura de poesías en voz alta. Lectura independiente Combinaciones de vocales Uso de adjetivos Supervisión de los pares Mostrar, no decir. Reflexión</p>	<p><i>Destrezas y procesos</i> Escribir cartas: carta amistosa. Discusiones: predicción. Lectura de poesía en voz alta. Apuntes de diferentes culturas. Ordenamiento en secuencias, relatos, intercambio de ideas.</p>	<p><i>Destrezas y procesos</i> Describir rasgos de los personajes. Reconocer adjetivos, verbos, sustantivos. Utilizar elementos de un cuento de hadas. Apuntes en la agenda diaria, reflexión. Trabajo cooperativo. Lectura independiente. Utilización de palabras compuestas.</p>
<p><i>Evaluaciones</i> Agendas diarias para anotar reflexiones sobre textos literarios. Orden alfabético para deletrear palabras. Localización de determinadas palabras. Escritura individual. Creación de diagrama propio. Ayudar al compañero con la lectura. Idear un diagrama creativo. Localizar y enumerar adjetivos en un texto literario.</p>	<p><i>Evaluaciones</i> Agendas diarias para anotar reflexiones sobre textos literarios. Orden alfabético para deletrear palabras. Localización de determinadas palabras. Escritura individual. Ayuda al compañero con la lectura. Escribir el borrador de una carta amistosa. Predecir y discutir aspectos relativos a las ballenas. Jugar diversos juegos de esparcimiento procedentes de diferentes culturas. Ordenar cuentos en secuencias y relatarlos. Intercambio de ideas acerca de lo que saben, quieren saber y han aprendido.</p>	<p><i>Evaluaciones</i> Agendas diarias para anotar reflexiones sobre textos literarios. Orden alfabético para deletrear palabras. Localización de determinadas palabras. Escritura individual. Escribir borradores de una carta amistosa. Predecir y discutir aspectos relativos a las ballenas. Intercambiar ideas acerca de lo que saben, quieren saber y han aprendido acerca de los pueblos nativos de América del Norte. Comparar /diferenciar elementos del cuento de hadas utilizando el diagrama de Venn. Compartir trabajos escritos con la clase para recibir retroalimentación y efectuar correcciones.</p>

Un proyecto de resolución de problemas matemáticos en cuarto grado

Warren Whitney, maestro de cuarto grado del distrito escolar North Shore, en Glen Head, Nueva York, centralizó su portafolio profesional en el diseño de tareas de evaluación relativas a la resolución de problemas en el área de las Matemáticas. Tras elaborar un

diagrama detallado de los resultados previstos y sus indicadores, Warren combinó las actividades de diseño de las evaluaciones con determinados resultados concretos. Estos se relacionaban con el desarrollo de la capacidad de los alumnos para resolver problemas, aplicar estrategias y presentar información matemática. El siguiente es un comentario extraído de su portafolio.

Presenté problemas nuevos a la clase cada semana. El trabajo de los alumnos en un problema se dividía en cuatro categorías. En la categoría comprensión, el alumno exploraba qué planteaba el problema, que información se requería y cual sería una respuesta razonable. La categoría explicación abarcaba todas las ideas, planes y ensayos requeridos para resolver el problema. A continuación, el alumno empujaba la presentación para explicar los métodos y las respuestas en forma verbal (la explicación) y en forma gráfica (la representación). Por último, en la categoría extensión, el alumno procuraba extraer conclusiones de sus respuestas y aplicarlas a situaciones generales.

Los alumnos abordarán un problema por semana y formularán su solución en borrador. Una vez por mes, seleccionarán uno de los problemas para “publicarlo. Elegirán, corregirán y pulirán los borradores. Las diez copias manuales corregidas se evaluarán sobre la base de una tabla de especificación que elaboraremos en clase y se guardarán en un portafolio de resolución de problemas. Las conversaciones entre los alumnos se concentrarán en determinar a qué parte de la tabla de especificación corresponde cada categoría de la resolución del problema “publicado”. Los alumnos autoevaluarán primero su trabajo, luego discutirán sus conclusiones conmigo y, por último, fijarán las metas para las siguientes actividades de resolución de problemas

En la figura 5.3 se muestra la tabla de especificación para la resolución de problemas matemáticos que ideó Warren.

Warren les indicó a sus alumnos que resolvieran este problema: *la señora López alquiló ocho docenas de sillas para una reunión. ¿Cuántas filas de sillas puede formar si pone ocho sillas en cada fila?*

Una alumna llamada Seema expresó en los siguientes términos su comprensión del problema:

Lo que sé (a partir del texto)

Se que el problema no dice qué forma tiene la habitación.

Lo que quiero saber (a partir del texto):

Quiero saber cuántas filas puede formar la señora si cada fila tiene ocho sillas.

Lo que quiero saber (por mí misma):

Quiero saber cuantas filas pueden formar si la habitación es redonda.

Supuesto

Supongo que no habrá ninguna diferencia.

La explicación de Seema para el problema fue esta:

La señora López puede tener 12 filas de sillas si pone 8 sillas en cada fila. Hay 12 columnas y 8 sillas, así que hice un rectángulo de 12x8. Hice un cuadrado para mostrar cómo resolver el problema, había 96 sillas en total, así que $8 \times 12 = 96$ y hay 96 sillas en mi cuadro.

La extensión de Seema para el problema de las sillas es:

Mi extensión es ¿cuántas filas más habría si la habitación fuera redonda y las sillas tuvieran distintos tamaños?

A continuación, Seema incluye la siguiente extensión y explicación:

En mi extensión volví a obtener 12 filas. Pensé que tendría la misma cifra que antes por que en tanto tenga 8 docenas de silla y haya 8 en cada fila, siempre se obtendrán 12 filas. También descubrí que para que queden bien, los espacios entre las sillas tienen que ser diferentes.

Los portfolios en una clase de sexto grado

Además de las evaluaciones del desempeño, los docentes suelen diseñar portafolios para que sus alumnos los usen. A medida que los docentes se familiarizan cada vez más con su empleo, el diseño del portafolio va pasando de ser totalmente decidido por el docente a estar rentado y pertenecer al alumno. Los portolios dejan de ser documentos para uso exclusivo de los docentes y adquieren destinatarios y finalidades reales, al margen de la asignación de calificaciones. También documentan e rendimiento, el esfuerzo y el progreso de los alumnos hacia el logro de resultados significativos y claramente establecidos. El lector puede ver fácilmente cómo se inserta el portafolio en el currículo y qué intervención tienen en su elaboración los docentes, los alumnos y otras personas. Estos portafolios deben incluir una descripción precisa del modo en que los alumnos seleccionan sus trabajos y reflexionan al respecto, así como de la manera en que el docente y los alumnos evalúan el portafolio.

Lisa Boreum, docente de Educación Especial, y Binnie Burke, maestra de sexto grado, ambas del distrito escolar Sag Harbor, en Sag Harborg, Nueva York, diseñaron un portafolio orientado al alumno para la clase de Bonnie. Su portfolio profesional se centraba en capacitar a los alumnos para que demostraran su progreso y su rendimiento en las actividades de investigar, pensar creativamente, leer y hacer presentaciones orales. Los alumnos, a su vez, seleccionaban dos de estos resultados y su rendimiento, los alumnos seleccionaban dios entradas casa una de estas dos áreas: una entrada inicial y un proyecto paralelo posterior. Durante el año, las dos docentes les proporcionaron a los alumnos una serie de herramientas de evaluación que ellos podían usar para evaluar a los alumnos identificar visualmente la relación entre determinadas tareas asignadas y los cuatro resultados del portafolio (veáse la figura 5.4).

Figura 5.3

Tabla de especificación para la resolución de problemas en y una clase de Matemáticas de cuarto grado

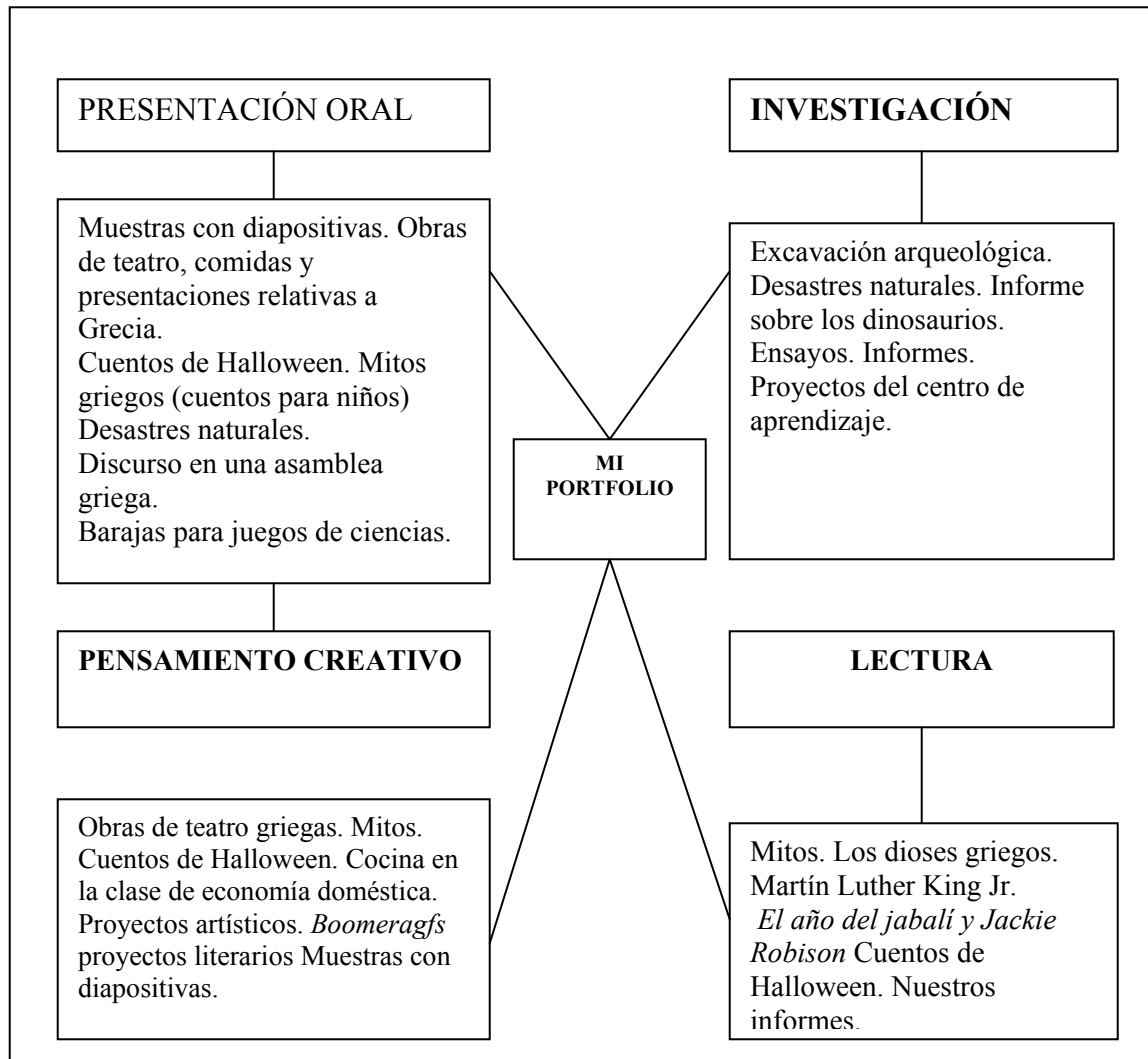
DIMENSIÓN	1	2	3	4
<i>Comprensión</i>	No entendió el problema.	Entendió lo bastante como para resolver parte del problema	Entendió todo el problema	Encontró factores especiales que afectan la resolución del problema.
<i>Explicación</i>	No dio ninguna explicación	Explicó el problema, pero el lector tuvo que completar algunos pasos.	Explicó todos los pasos de la resolución	La explicación contenía información adicional sobre lo que estaba pensando el encargado de resolver el problema.
<i>Representación</i>	No hizo ninguna representación	El lector tuvo que llenar algunos detalles que faltaban para entender la representación.	La representación fue clara, organizada y bien rotulada.	La representación fue esmerada (se utilizaron una o dos representaciones completamente diferente):
<i>Extensión</i>	No hizo ninguna extensión	Hizo algunos comentarios sobre el problema.	Aplicó la solución a otra área del problema.	La extensión llegó a una regla general (o conclusión) acerca del problema.
<i>Representación global</i>	La copia final estaba desprolija y desorganizada	Se tomó el cuidado de hacer una copia final prolija y organizada	La copia estaba bien escrita, prolija y colorida.	La copia final estaba prolija, bien escrita y contenía elementos especiales.

También se les suministró a los alumnos una lista de proyectos que servirían para demostrar los diferentes resultados, como por ejemplo, proyectos literarios, discursos, mapas en relieve, caricaturas políticas, representaciones teatrales, test, afiches y lecciones. En sus presentaciones, los alumnos debe explicar de qué manera el proyecto da prueba de su capacidad de aprendizaje en una determinada categoría (investigación, lectura, presentación oral o pensamiento creativo). Los alumnos pueden usar proyectos correspondientes a cualquiera de sus clases; estos son solo unos pocos ejemplos.

Se incluían otras herramientas, como auto evaluaciones y preguntas para la reflexión, que los alumnos debían completar regularmente. Los siguientes son algunos ejemplos de esas preguntas:

- ¿cómo has usado las pautas del proyecto para ayudarte a realizar las tareas requeridas de la mejor forma posible?
- ¿cómo has empleado tus propias ideas para realizar la tarea de la mejor forma posible?
- ¿Cómo te han ayudado tus actitudes y tus hábitos de trabajo?
- ¿Cómo te han impedido tus actitudes y hábitos de trabajo realizar esta tarea de la mejor forma posible?
- ¿Qué decisiones has tomado que te han creado obstáculos para realizar tu proyecto?
¿Cómo superaste esos obstáculos?
- ¿Qué métodos de recolección de datos te resultaron más útiles?
- ¿Cómo has ayudado a otros alumnos a responder preguntas, resolver problemas, etcétera?

Figura 5.4
Una red de proyectos elaborada por una clase de sexto grado.



Para cada entrada del portfolio, los alumnos suministran información detallada sobre el contexto. En la figura 5.5 se enumeran todos los requisitos que deben cumplir los alumnos al seleccionar una entrada para su portfolio. La figura 5.6 muestra una lista de preguntas sobre las cuales los alumnos deben reflexionar después de haber seleccionado proyectos para sus portfolios.

Como puede observarse en las pautas de la figura 5.5, los alumnos de la clase de Lisa y Bonnie asumen la responsabilidad de comentar y explicar los trabajos incluidos en sus portfolios. Con esa información, cualquier lector puede apreciar la cantidad de esfuerzo, tiempo y respaldo que recibió cada alumno para producir los trabajos incluidos.

En su etapa más avanzada, las evaluaciones incluidas en los portfolios del docente y de los alumnos son verdaderamente auténticas. Esto quiere decir que tienen los siguientes atributos:

- Son tareas sustanciales, insertas en el currículo, que les exigen a los alumnos utilizar conocimientos previos, aplicar conocimientos y destrezas correspondientes a una o más áreas de contenidos, expresar detalladamente sus conclusiones, reflexionar y revisar su trabajo.

- Son valorados por destinatarios externos a la escuela, tiene un propósito real para los alumnos y los docentes, y son lo suficientemente flexibles como para darles opciones y oportunidades de éxito a todos los alumnos.
- Toman en cuenta diferentes necesidades y orígenes culturales.
- Los correspondientes estándares de desempeño están claros para todos los interesados. El docente y los alumnos los identificaron conjuntamente y los articularon en las tablas de especificación, los criterios de puntuación y los ejemplares. Estos estándares sirven de guía a los alumnos para evaluar su trabajo y fijarse metas para mejorarlo.

Figura 5.5

Pautas para alumnos de sexto grado respecto de la selección de entradas de portfolio.

Recursos

Para ayudarte a evaluar tu trabajo con mayor claridad, hemos adjuntado a casa sección:

1. Una lista de destrezas o cualidades que pueden empear los alumnos de sexto grado en una diversidad de proyectos.
2. Una lista de proyectos posibles correspondientes al año lectivo entre los cuales puedes elegir las piezas para tu portfolio.

Procedimiento

1. Lee cuidadosamente las listas de destrezas y proyectos.
2. Decide qué destrezas o cualidades se ponen en evidencia en cada uno de posproyectos que has seleccionado.
 - a. Haz de las destrezas evidenciadas en cada proyecto.
 - b. Decide si la evidencia ilustra un punto fuerte, un punto débil o algo que está mejorando
 - c. Explica cómo o porqué sabes esto.
3. Reúnete con las señoritas Burke o Boerum para comentar qué trabajos has seleccionado y qué pruebas suministran de tu aprendizaje.
4. De esos comentarios, toma las pautas de reflexión apropiadas que te ayudarán a explicar tus puntos fuertes y débiles y el modo en que crees haber mejorado durante el año.
5. Tómate tiempo para responder con cuidado y en oraciones completas las preguntas de reflexión.
6. Revisa tus respuestas con las señoritas Burke o Boerum para ver si has dado prueba de sus destrezas, procesos y progreso intelectuales.
7. Guarda en tu portfolio las piezas seleccionadas, las tablas de especificación, las listas de destrezas y las auto evaluaciones

Puedes modificar la selección de trabajos que has hecho para el portfolio en cualquier momento en que consideres que otro proyecto da mejor prueba de tus destrezas para aprender.

Creado por L: Boreum y B. Burke. Corregido por L. Ruhl, reproducción autorizada

tabla de especificación y usar los niveles indicados para consolidar su aprendizaje y establecer metas concretas. Los niveles más bajos superan en una excelente herramienta de andamiaje y de enseñanza. El nivel superior está por encima del estándar esperado, pues hasta el alumno de mejor rendimiento es alentado a mejorar. Además, la tabla de especificación incluye ejemplos reales para cada nivel, que los alumnos ayudaron a identificar y evaluar.

A continuación, veremos algunos ejemplos de tareas auténticas, tablas de especificación y reflexiones que he seleccionado de los portfolios de diversos docentes

Proyecto relativo al estudio de un libro en una clase de lengua de sexto grado

Nuestro siguiente ejemplo es un proyecto relativo al estudio de un libro, tomado del portfolio de Lisa Boreum y Bonnie Burk Estas son las instrucciones que se les dieron a los alumnos:

Todos los alumnos deben realizar al menos un proyecto relativo al libro que están leyendo. Pueden trabajar junto con otros alumnos que hayan leído el mismo libro, pero todos serán responsables de las opciones que hagan. Se prevé que el proyecto tendrá mayor calidad si dos o más alumnos trabajan juntos. Los alumnos que ayuden a otros a realizar su proyecto obtendrán puntos adicionales por el esfuerzo de prestar esa antes de completar el proyecto. Una vez terminado este, utilizarán la tabla de especificación para evaluarlo.

De los proyectos aquí sugeridos, el alumno debe elegir uno para realizarlo. Si se le ocurre otro proyecto alternativo, la señorita Burke debe aprobarlo antes de que el alumno lo inicie.

Figura 5.6

Preguntas para guiar las reflexiones de los alumnos de sexto grado sobre su selección de trabajos a incluir en el portfolio.

Una vez que has seleccionado los proyectos a incluir en tu portfolio, analizarás tu proceso de aprendizaje utilizando las preguntas de reflexión aquí indicadas. Estas preguntas te ayudarán a explicar:

1. Cómo procediste a elaborar el trabajo en cuestión.
2. Cuáles son tus puntos fuertes y tus puntos débiles.
3. Qué te resultó más fácil y qué te resultó difícil de la realización del proyecto.
4. Cómo revisaste tu trabajo.
5. Cómo podrías mejorar tu trabajo.
6. En qué has mejorado desde comienzos del año, junto con datos en tus reflexiones que lo demuestren.
7. Cómo crees que ha mejorado tu capacidad de reflexionar desde comienzos del año, junto con pruebas que lo demuestren.
8. Cómo te parece que trabajó tu grupo en el proyecto, y cómo crees que el trabajo podría haber sido mejor.
9. Qué recursos utilizaste para realizar el proyecto.
10. Por qué seleccionaste los proyectos que incluiste en tu portfolio.
11. Qué dicen los trabajos del portfolio acerca de tu forma de aprender.

No necesitas limitarte a estas cuestiones. Puedes agregar cualquier opinión que desees acerca de tus trabajos: qué te gusto o no te gustó, qué te complace o qué te desagrada respecto de la realización de los proyectos.

Puede ser útil que vuelvas a leer las reflexiones que ya has escrito en los trabajos. Luego podrás comentar si te parece que has explicado correctamente las destrezas que utilizaste y los procesos que recorriste para realizar el proyecto. También podrás explicar como han aumentado tus expectativas sobre tu propio aprendizaje. ¿Tienes más capacidad para aprender ahora? De ser así ¿qué pruebas puedes aportar para demostrar el incremento de tu capacidad?

Cada sección de tu portfolio debe incluir:

1. Proyectos seleccionados que den prueba de tu capacidad emergente en desarrollo de aprender en cada área.
2. las tablas de especificación para el proyecto.
3. Una descripción de cada proyecto, sus partes desde el comienzo hasta el final y una explicación de tu rol en el proyecto.
4. las destrezas que se ponen de manifiesto en cada proyecto:
 - a. ¿Cuáles son sobresalientes? ¿Cómo puedes determinarlo?
 - b. ¿Cuáles requieren más trabajo? ¿Cómo puedes determinarlo?
 - c. ¿Cuáles han mejorado? ¿Cómo puedes determinarlo?
5. Todas las autorreflexiones que escribiste acerca de cada proyecto: tu proceso, tu desempeño, tus metas para mejorar tus destrezas y tus estrategias para cumplir o alcanzar esas metas en proyectos futuros.
6. Todos los materiales aportados por ti y por tu maestra que tengan relación con los proyectos seleccionados para cada sección de tu portfolio.

Actuación en vivo o grabada en video

(Segmento de 3 a 10 minutos de duración)

- ♣ *Programa en vivo.* Conducirás un programa con invitados que han trabajado en el libro o son expertos en el tema central del relato. puedes pedirle al público que hable de sus experiencias al respecto.
- ♣ *Escena.* Representarás una escena del libro que muestre el punto crucial del argumento. Harás una introducción para poner el argumento en conocimiento del público.
- ♣ *Programa informativo.* Harás un programa informativo sobre el proceso de escritura del libro. Recuerda incluir datos acerca de la investigación requerida para escribir el libro.
- ♣ *Cuento* si estás leyendo un a serie de cuentos, escribe y representa una escena que incluya a los diferentes personajes como su se hubieran encontrado. Todos deberán conservar las características que tienen en sus respectivos cuentos.
- ♣ *Video de danza o música.* Realizarás un video de danza o música que muestre la trama del relato, los conflictos, el punto crucial y la resolución. No emplees palabras.
- ♣ *Canción.* Compondrás o ejecutarás una canción acerca del argumento del libro.

Dibujo, pintura o escultura

(Tendrás que presentar tu proyecto y defenderlo ante la clase).

- ~ Crearás una obra artística que ilustre la trama de la historia y su resolución.
- ~ Crearás una obra artística que muestre los principales personajes. Debes incluir información sobre las características de los personajes.
- ~ Harás un juego de mesa que incluya elementos de la historia. No es necesario que los jugadores hayan leído el libro para participar en el juego. Recuerda incluir los correspondientes elementos e instrucciones.

Escritura

- ~ Escribirás e ilustrarás un libro infantil en el que se cuente la historia a un público de menor edad. Tómate tu tiempo para hacer las ilustraciones, y considera el medio con el que estarás trabajando.
- ~ Armarás un libro de recortes de las notas y apuntes de investigación que el autor podría haber reunido para escribir el libro. Incluye cinco entradas en una agenda diaria que el autor podría haber llevado en el proceso d escribir la historia.
- ~ Escribirás un juego de computadora e instalarás el programa de modo que sea compatible con una de las computadoras de la escuela. El juego requerirá que los participantes respondan preguntas sobre los elementos literarios del libro.
- ~ Cualquiera sea el proyecto que elijas, ¡no olvides leer. La tabla de especificación!

En la figura 5.7 se muestra una plantilla de calificación para el proyecto literario descripto. El docente y los alumnos van llenando la planilla a medida que trabajan en el proyecto. El docente utiliza la tabla de especificación (figura 5.8) para volcar las puntuaciones en una calificación.

A diferencia de lo que ocurre con las tareas normalmente asignadas para el estudio de un libro, Lisa y Bonnie toman en cuenta la importancia de crear productos que tengan valor

en el mundo exterior. Le dan a los alumnos la oportunidad de emplear sus estilos de aprendizaje preferidos y su inteligencia.

Un viaje por el hemisferio occidental en una clase de quinto grado

Otro ejemplo del uso de la reflexión y las evaluaciones contextulizadas proviene del portfolio de Kristinna Marinadle, maestra de quinto, que emplea una evaluación diagnóstica interdisciplinaria de Geografía y Lengua con sus alumnos, la evaluación fue elaborada junto con otro docente d quinto grado, Stu Dick, y utilizada en las clases de Kristinna y de Stu. Las siguientes son las instrucciones para la tarea asignada:

En las últimas semanas, nuestras dos clases dedicaron mucho tiempo a la elaboración de este interesante proyecto. Ustedes ya pensaron prácticamente en todo. De hecho, ¡nos parece que sus ideas son mejores que las nuestras! Aportaron una cantidad de ideas distintas y excelentes. Ahora van a trabajar en equipo, van a investigar mucho, van a aplicar todos los pasos del proceso de escribir y muchas cosas más ¡esperamos que todos trabajen al máximo de sus posibilidades en este proyecto!

Estos son los requisitos que plantearon:

1. Tus compañeros y tú escribirán un cuento sobre un viaje desde un lugar del hemisferio occidental a otro.
2. Tus compañeros y tú elegirán un lugar como punto de partida y otro como destino final. La distancia de un lugar al otro será de 2.000 millas o más. Uno de nosotros deberá aprobar la selección de los lugares.
3. El cuento puede tener personajes ficticios, pero todo lo que ellos vean o encuentren debe basarse en hechos que ustedes hayan averiguado en su investigación. En otras palabras, su cuento debe ser realista.
4. Como en todas las buenas historias, los personajes deben sortear algún problema.
5. Los miembros de tu equipo deben elegir una época para le viaje, de modo que sepan con qué tipo de personas se van a encontrar y describan cómo viven.
6. El equipo debe utilizar un mínimo de tres medios de transporte durante el viaje.
7. El equipo debe encontrarse con un mínimo de dos tipos de pueblos indígenas.
8. El equipo debe discutir los entornos de los personajes, incluyendo aspectos geográficos, climáticos, etcétera. Deberán investigar mucho.
9. El equipo trazará un mapa del viaje.
10. El equipo hará anotaciones en una agenda diaria al final de cada sesión de trabajo. En esta agenda, tus compañeros y tú escribirán lo que han aprendido y logrado en cada sesión. También reflexionarán sobre el modo en que han trabajado en equipo y lo que debe hacer para mejorar ese trabajo.

Si quieren hacer algo que no está contemplado en estos requisitos, hablen con uno de nosotros y analizaremos si podemos aprobar su idea. Por favor, tengan presente que este proyecto se les ocurrió a ustedes. A nosotros nos parece una excelente oportunidad para que sean todo lo creativos que puedan y disfruten de esta interesante parte del mundo. ¡Sabemos que va a gustarles!

Afectuosamente
Srta. Martindale y Sr. Dick

Las tablas de especificación y las preguntas de reflexión para el proyecto del hemisferio occidental aparecen en las figuras 5.9 y 5.10.

No siempre es fácil crear todas las condiciones propicias para que los alumnos perciban el significado y el valor que puede tener su trabajo fuera del aula. Algunos docentes superan esta dificultad proponiendo tareas cuyo significado fuera de la escuela es evidente, o creando en el aula condiciones semejantes a las de la vida real.

Figura 5.7

Alumno	docente	
_____	_____	1 =El alumno entregó una propuesta en la que se describe un plan del proceso requerido para completar el proyecto – adjuntando información sobre el tiempo, los materiales y los recursos necesarios- e incluyó una predicción sobre su aprendizaje.-
_____	_____	2 Durante el transcurso del proyecto, el alumno demostró comprender el proceso de investigación que efectuó el autor para realizar el libro.
_____	_____	3 El proyecto fue apropiado para el libro.
_____	_____	4 El alumno presento el proyecto de una manera clara.
_____	_____	5 El proyecto fue prolijo, organizado y claro. El alumno aprovechó el medio con el que trabajo, utilizándolo en la forma indicada.
_____	_____	6 El alumno evalúa atentamente su propio proyecto, empleando ejemplos para señalar sus puntos fuertes y débiles.
_____	_____	7 El alumno tuvo en cuenta al público y siguió las instrucciones de la tarea asignada.
_____	_____	8 Durante el transcurso del proyecto, el alumno demostró comprender los elementos literarios del libro (argumento, tema, personajes, ambiente, conflicto, punto crucial, resolución, etcétera).
_____	_____	9 El alumno manejó su tiempo de modo de cumplir los plazos del proyecto.
_____	_____	10 El proyecto demostró creatividad.
Instrucciones: califica tu proyecto y el proceso de realizarlo. Lee atentamente cada objetivo. Asígna una calificación de 5 a 0 (5 = Excelente; 0= Mal).		
Comentarios		

Plantilla de calificaciones para proyectos literarios en una clase de Lengua de sexto grado

Figura 5.8
Tabla de especificación para proyectos literarios
en una clase de Lengua de sexto grado

<p>A + El alumno:</p> <ul style="list-style-type: none"> * Entregó una propuesta detallada en la que se describen el proceso, los materiales y los recursos necesarios para realizar el proyecto. * Realizó un proyecto apropiado para el libro en cuestión. * Demostró durante el proyecto su comprensión de la investigación efectuada y la información relativa a los elementos literarios del libro (argumento, conflicto, complicaciones, punto crucial, resolución, tema, personajes, etcétera). * Manejó su tiempo con eficacia haciendo arreglos de antemano para el trabajo (asociaciones, materiales, lugar, recursos, tiempo, etcétera). * Tuvo en cuenta al destinatario del proyecto. * Realizó un proyecto final prolijo, organizado y claro. * Presentó el proyecto de manera clara. * Evaluó atentamente su propio trabajo por vía de describir sus puntos fuertes y débiles. <p>A Proyecto muy similar al de A+. Reflejó la comprensión del alumno de los elementos literarios del libro y del proceso de su realización. El alumno lo completó de un modo eficaz.</p> <p>B Este proyecto estuvo muy bien hecho. El alumno:</p> <ul style="list-style-type: none"> ^ Evaluó su propio trabajo utilizando ejemplos para señalar sus puntos fuertes y débiles. ^ Podría haber desarrollado más algunos elementos del proyecto. ^ Tal vez no haya comprendido todos los elementos del libro o cómo usar correctamente el medio. <p>C Este proyecto tuvo algunas cosas bien hechas pero su calidad global fue insatisfactoria. El alumno:</p> <ul style="list-style-type: none"> ^ No utilizó el proyecto para comunicar los elementos literarios del libro no el proceso de realización. ^ Entregó un proyecto que no se adecuaba al libro en cuestión. ^ Omitió entregar una propuesta detallada del proyecto. ^ No cumplió los plazos de entrega del proyecto. ^ Eligió trabajar con personas que no podían reunirse o no trabajan bien juntas. ^ Entregó un producto final desprolijio, desorganizado y confuso. ^ Presentó el proyecto de una manera apresurada y poco clara. ^ No pudo señalar los puntos fuertes y débiles del proyecto. <p>N Este proyecto no siguió las pautas de la tarea asignada o no fue realizado.</p>

Figura 5.9

Tabla de especificación para el proyecto de un viaje

4=	El viaje es acertado en sus aspectos histórico y geográfico, y guarda fidelidad con la época elegida. Es evidente que se hizo una profunda investigación, la que está creativamente insertada en el relato. El viaje sigue una línea narrativa lógica y es entretenido e interesante de leer o escuchar.
3=	La mayor parte de la información geográfica, histórica y de época es correcta. La investigación previa realizada se pone de manifiesto en el relato y está claramente transmitida. El viaje sigue en una línea narrativa lógica.
2=	El viaje sigue una línea narrativa lógica. Parte de la investigación previa no está correctamente transmitida o es poco clara. O: la mayor parte de la investigación es correcta y está claramente transmitida, pero la línea narrativa es ilógica y difícil de seguir.
1=	El viaje contiene graves errores en la relación entre la información histórica, geográfica o de época. Se ha efectuado escasa o ninguna investigación previa. La línea narrativa es imposible de seguir y el relato es poco interesante de leer o escuchar.

por el hemisferio occidental

figura 5.10

Preguntas de reflexión para el proyecto de un viaje por el hemisferio occidental.

Nombre: _____

Fecha: _____

Nombre del grupo: _____

Reflexiona sobre la sesión de trabajo de hoy y responde estas preguntas en la forma más completa que te sea posible:

1. ¿qué lograste hoy? (Incluye cualquier avance en tu proyecto o cualquier cosa que hayas aprendido hoy)

2. ¿Qué vas a hacer tú, individualmente, durante la próxima sesión de trabajo?

3. ¿Qué ejemplos de un buen trabajo en equipo pusieron de manifiesto tu grupo en el día de hoy? (Puedes incluir cómo resolvieron problemas juntos, qué hicieron cuando no estaban de acuerdo, cómo se trataron los miembros del equipo unos a otros, etcétera).

4. ¿qué tiene que mejorar tu equipo para la próxima sesión y cuál es tu plan de acción para efectuar esta mejora?

Proyecto de un aviso de prensa en una escuela primaria

Joanne Picote-Zocchia, especialista en Lengua del nivel primario en el distrito escolar de West Islip, Nueva York, elaboró y asignó una tarea interdisciplinaria de matemáticas y lengua utilizando la sección de avisos clasificados de un diario. La tarea asignada a sus alumnos de sexto grado fue:

Utiliza la sección de avisos clasificados de un diario para hacer lo siguiente:

1. ¡Te has quedado sin empleo! Ve a la sección de avisos clasificados y encuentra otro. Recorta el aviso correspondiente y pégalo en el espacio provisto. Recuerda que debes ganar lo bastante como para mantenerte. Calcula tus salarios semanales y mensuales.
 -Salario Semanal: _____
 -Salario mensual: _____
2. tu nuevo empleo te obliga a mudarte a otro lugar. Decides alquilar transitoriamente una vivienda, hasta que conozcas mejor la zona. Encuentra un departamento, casa o habitación para alquilar (resta el precio del alquiler de tu salario mensual para asegurarte de poder pagarlo- muestra tus cálculos). En el espacio de abajo, pega el recorte del aviso de tu nuevo hogar:
 _____ Salario mensual
 _____ Alquiler mensual

3. Vas a necesitar un auto para ir y volver del trabajo. Encuentra un automóvil confiable y de precio razonable. Para calcular las cuotas mensuales que pagarás, divide el precio del auto por 36. resta la cantidad resultante de la suma mensual que te queda después de pagar el alquiler. En el espacio de abajo, pega el recorte del aviso del auto y muestra todos tus cálculos.

Tras completar el proyecto, los alumnos contestan las siguientes preguntas:

- ✧ ¿Cómo te sentiste mientras hacías esta tarea?
- ✧ ¿Qué parte de la tarea te resultó más difícil?, ¿más fácil?, ¿más divertida?
- ✧ Si pudieras volver a hacer esta tarea ¿qué harías de distinta manera?
- ✧ ¿Cuáles son las cosas más importantes que aprendiste?

En la figura 5.11 se muestra la tabla de especificación para este proyecto, que fue elaborada por los alumnos de Joanne a los efectos de evaluar el resultado de su proyecto de encontrar empleo y presupuestar sus recursos.

Presentaciones en una clase de séptimo grado.

Linda Hughs es maestra de séptimo grado en el distrito escolar de Manhasser, Nueva York. Linda procuró hacer más significativas y auténticas las presentaciones de los alumnos instándolos a que crearán una situación simulada para presentar trabajos de investigación basados en un mínimo de tres obras literarias. Los alumnos debían terminar por lo menos tres destinatarios distintos que fueran apropiados para su presentación,

seleccionar uno de los tres y luego diseñar una presentación de seis minutos en la que adoptarían una personalidad específicamente adecuada para ese público. Se les dieron las siguientes instrucciones:

Primero, te sentarás en el pupitre del lado izquierdo del aula. En ese pupitre, *serás tú mismo* y nos explicarás qué has creado para tu presentación.

1. nos dirás qué rol vas a adoptar como presentador de esta información que vamos a escuchar. [por ejemplo: *mi nombre es Linda Hughs y estoy cursando mi primer año en la universidad Kenyon, donde estudio biología. El año pasado me ofrecieron ser directora estudiantil de la Casa de Arácnidos de la Universidad. Hoy he venido a dar una charla en el Filial de Manaste de la Sociedad de Arácnidos*].
2. luego definirás el auditorio ante el que piensas hacer tu presentación. (los que te escuchemos tendremos entonces que pensar como lo haría ese auditorio) [*la sociedad de Arácnidos es un grupo de adolescentes interesados en las arañas y otros insectos especulantes. Esta reunión es la primera de una serie de charlas a cargo de distintos oradores invitados para tratar la importancia de la biología en la educación universitaria.*].
3. luego nos harás conocer el propósito de tu presentación. [*Mi propósito es contarles mi experiencia con las arañas con la que trabajé el año pasado y cómo se despertó mi interés por estas fascinantes criaturas.*]

A continuación, irás al lugar del aula en el que vas a hacer tu presentación. Si tienes que cambiar de lugar los muebles, por favor házmelo saber el día anterior para poder ayudarte. Te pondrás el vestuario para la presentación al principio de la clase. Asegúrate de que los elementos visuales (si vas a usarlos) estén listos.

Tu presentación debe durar unos seis minutos. (lamentablemente, estamos un poco cortos de tiempo, así que debes estar listo para empezar en cuanto sea tu turno.)

Por último, cuando hayas terminado, debes guardar todos tus elementos mientras el siguiente alumno se prepara para hacer su presentación.

Nota: habrá un estrado disponible, y también un par de pupitres que puedes usar como mesa o escritorio. Además, pienso grabar en video algunas de estas presentaciones y posiblemente necesite contar con voluntarios.

La tabla de especificación para esta tarea aparece en la figura 5.12 la figura 5.13 muestra un ejemplo del esquema de una presentación que hizo un alumno de séptimo grado.

TRABAJOS ESCOLARES

Como muestran los ejemplos que he presentado, las evaluaciones incluyen consignas que orientan a los alumnos a reflexionar sobre los productos y los procesos, y que pueden tomar la forma de preguntas concretas, listas de verificación o tablas de especificación. Las reflexiones e indicaciones demuestran que el docente entiende la necesidad de adecuarse a diferentes niveles de desarrollo y estilos de aprendizaje. la esencia del portfolio de un alumno es la reflexión que él hace sobre sus trabajos.

El primer ejemplo al respecto es un registro semanal de reflexiones que Andrea Gerstenblatt utiliza con sus alumnos de segundo grado.

Figura 5.11

Tablas de especificación para el proyecto de encontrar un empleo por el diario

	Experto (4)	Avanzado (3)	Novato (2)	Principiante (1)	¡ah, ay, ah...! (0)
Computación	Sin errores	El método elegido es correcto pero hay dos o tres errores en los cálculos	El método elegido es incorrecto, o hay más de tres errores en los cálculos	El método elegido es incorrecto o hay errores en los cálculos	el método elegido es incorrecto y hay errores en computación
Cumplimiento	El trabajo esta completo según las instrucciones	Faltan títulos	Faltan avisos	Faltan cálculos	Faltan avisos y cálculos
Organización	Todo está donde corresponde	El trabajo se sale del espacio asignado	El trabajo de un espacio interfiere con otro	El trabajo está en el espacio correcto	El trabajo no parece tener ninguna organización.
Prolijidad	Sin errores ni correcciones	Claro y fácil de leer.	Hay cambios y/o correcciones que distraen al lector.	Hay cambios y/o correcciones que dificultan que el lector pueda entender las respuestas	El trabajo es ilegible
Puntualidad	puntualmente	Puntualmente realizado pero no entregado	Terminado el día de entrega y entregado antes de finalizar el día	Un día de atraso	Más de un día de atraso
Proceso	Utilizado por el docente para ayudar a otros alumnos	Trabajo completado sin ayuda	Trabajo completado con ayuda de los pares	Trabajo completado con ayuda e un adulto	No solicitó la ayuda necesaria

Figura 5.12
Tabla de especificación diseñada por docente y alumnos para un proyecto de presentación de séptimo grado

Alumno: Evaluador:				
	Ejemplar	Excelente a Muy bueno	Bueno a Necesita mejorar	Insuficiente a inaceptable
Creación del público	<ul style="list-style-type: none"> describe vívidamente el auditorio. Hay una perfecta adecuación entre el auditorio y la presentación. El auditorio descriptivo evidentemente existe en el mundo real 	<ul style="list-style-type: none"> Describe un auditorio concreto. Hay una clara adecuación entre el auditorio y la presentación. El auditorio descriptivo, podría existir en el mundo real. 	<ul style="list-style-type: none"> Describe al auditorio en términos generales o vagos. Hay cierta adecuación entre el auditorio descriptivo pero en forma excepcional o forzada. 	<ul style="list-style-type: none"> No se describe el auditorio. No hay adecuación entre el público y la presentación. El auditorio en cuestión no existe en el mundo real.
El rol de presentador	<ul style="list-style-type: none"> Representas perfectamente el rol; tu actuación es digna del “oscar”. El auditorio se olvida de que estás representando un papel. Usas vestuario y utilería auténticos Tu punto de vista parece indiscutible 	<ul style="list-style-type: none"> El rol es creíble en la presentación. Te mantienes “en el personaje” durante toda la presentación. El vestuario y la utilería son apropiados para el rol. Tu punto de vista es convincente y atinado. 	<ul style="list-style-type: none"> El rol se manifiesta escasamente. Te “sales del personaje” ocasionalmente. El vestuario y la utilería no son adecuados para el papel. Tu punto de vista es escasamente atinado. 	<ul style="list-style-type: none"> El rol no pone de manifiesto. No hay caracterización. No hay vestuario ni utilería. El auditorio no entiende tu punto de vista.
Contenido de la presentación y utilización de libros	<ul style="list-style-type: none"> La presentación responde exactamente a lo que se propuso y es del todo realista. El formato fue planeado en detalle. Hay referencias originales a los libros. Utiliza muchos libros. Las citas fueron esenciales para la estructura y en mensaje de la presentación. Pronuncia correctamente los títulos y los nombres de los autores y hace una variedad de referencias a los autores. 	<ul style="list-style-type: none"> Clara coincidencia entre el propósito de la presentación y su representación concreta. El formato tiene una estructura clara. Hay referencias a los libros en la presentación. Utiliza tres o cuatro libros. Las citas reafirman las ideas planeadas; hay un mínimo de tres citas. Pronuncia correctamente los títulos y los nombres de los autores. 	<ul style="list-style-type: none"> Es difícil encontrar la coincidencia entre el propósito de la presentación y su representación concreta Falta organización en el formato. Menciona libros pero no establece una relación clara. Menciona menos de tres títulos. Las citas tienen escasa conexión con el propósito y el tema de la presentación. Errores de pronunciación. 	<ul style="list-style-type: none"> No hay coincidencia entre el propósito propuesto y la presentación misma. El formato es desorganizado. No menciona ningún libro no hace ninguna referencia. Usa un solo libro. No hace citas. Pronunciación incorrecta.

En la figura 5.14 se muestran fragmentos de las reflexiones de una alumna llamada Jeanne y las respuestas de la maestra. Estos fragmentos muestran cómo fomenta Andrea la habilidad de los alumnos para establecer y monitorear sus metas de aprendizaje.

En el segundo ejemplo, que proviene de la clase de séptimo grado de Linda Hughs, es el resumen de las reflexiones de la alumna Sara sobre la presentación que hizo ante sus pares:

Sin duda aprendí mucho al crear esta presentación. Aprendí cómo actuar (es decir, aprendí diferentes características de las personas). Descubrí distintos tipos de vestuarios y comprobé cómo lo ayudan a uno a crear su papel. Por ejemplo, si uno quiere representar a un hombre de negocios, no puede aparecer vistiendo una remera y pantalones vaqueros. Los hombres de negocios generalmente usan traje y corbata. También aprendí otras características de la gente, como la manera de caminar. Un hombre de negocios suele caminar rápidamente, sin detenerse a observar nada. Además, aprendí un poco sobre cómo hacer grabaciones en video y películas. Escribir un guión no siempre es fácil. También hay que practicar mucho a los personajes antes de grabarlos.

Comprobé que pueden aprenderse muchas cosas reales en los libros de ficción. Por ejemplo, yo me entere de que el uso de drogas es el resultado de los conflictos familiares. Puede establecer conexiones entre mis siete libros y encontrar que el futuro de una persona con las drogas depende de su pasado con la familia. Casi todos los personajes de los libros habían tenido una mala influencia familiar, pero algunos los habían prevenido en contra de las drogas. También hacía otros elementos más útiles "ocultos" en mis libros. Por ejemplo, yo estaba leyendo sobre las drogas y mis libros traían información acerca de los distintos tipos de drogas y sus efectos colaterales.

Al llevar el registro del aprendizaje, aprendí a analizar y relacionar libros. Todas las noches, mientras leía pensaba en este libro en comparación con tal otro, anoté estas ideas en mi registro.

Otra cosa que aprendí al crear esta presentación es a SER CONCRETO. Sé que si no hubiera descripto a mi auditorio en detalle, no habrían tenido idea de lo que se estaba hablando. Comprobé que cuando algunos chicos utilizaban citas para describir sus libros, y podría hacerme un cuadro mental más claro de lo que ellos estaban tratando de enseñarme.

Creo que lo más importante que aprendí al crear esta presentación fue que trabajar con empeño tiene grandes recompensas.

El último ejemplo también procede de la clase de Linda Hughs, y es la carta al lector del portfolio que escribió una de sus alumnas:

Querido lector:

Me llamo Lauriana. Este cuaderno es la prueba absoluta de todo lo que hice durante esta investigación. Aquí se ponen en evidencia mi progreso y mi aprendizaje. Pasé, con sorprendente rapidez, de no tener mayor idea sobre el tema a adquirir un fuerte interés y un conocimiento cabal de la anorexia nerviosa. Aprendí muchísimo con esta investigación y el tiempo que le dediqué estuvo bien invertido. Trabajé con gran empeño en esto y me siento muy orgullosa. Los libros que leí sobre la anorexia nerviosa me resultaron interesantísimos y contundentes. Al principio tenía algunas dudas sobre el tema a elegir. Deseche varios libros y cambié de tema. Esto me llevo bastante tiempo, ahora que lo pienso. ¡Perdí mucho tiempo! El año pasado vi un artículo sobre la anorexia en una revista y realmente me interesó. Déjenme explicarles lo del registro del aprendizaje. Este registro es una agenda en la que fui anotando las cosas que hacía cada día. Allí se puede observar cómo fue surgiendo mi proyecto, en forma lenta pero segura. Los libros que leí

fueron fáciles de encontrar porque mi madre e los consiguió. Fue a la biblioteca y sacó cuatro libros que me encantaron. La presentación fue una gran experiencia de aprendizaje. Tuve que pararme frente aun público que yo había creado y enseñarles acerca de la anorexia. El trabajo fue arduo pero en su mayor parte valió la pena, sin ninguna duda.

Atentamente.

Lauriana

Figura 5.13
Esquema para un proyecto de presentación
de un alumno de séptimo grado.

Describe a tu auditorio:	Futuro(s) progenitore(s) Mujeres embarazadas
¿Qué quieren enseñarle a este auditorio?	Quiero enseñarle a este auditorio a disciplinar a su(s) hijo(s) sin golpearlos. También quiero enseñarles que el maltrato infantil es muy malo.
¿Qué formato usarás para enseñarle a este auditorio?	Usaría videos para mostrar los efectos en la vida de un niño maltratado y los severos castigos aplicados a quien (es) los maltrata(n).
¿Qué materiales necesitas?	Voy a necesitar un televisor y una cinta de video y un puntero (creo).

Para esclarecer el empleo de la evaluación y las reflexiones, Linda incluye ejemplos de trabajos de los alumnos junto con explicaciones de cómo funcionan las evaluaciones con diferentes alumnos. También señala el grado en que los trabajos escolares incluidos en su portfolio reflejan la diversidad de la clase, permitiéndole al lector extraer conclusiones acerca del valor del diagnóstico para una amplia gama de niveles de capacidad y conocimientos entre los alumnos.

Figura 5.14
Fragmentos del registro semanal de la alumna Jeanne, de segundo,
y las respuestas de su maestra.

1-16

Esta semana aprendí a trabajar con el portafolio.

Tuve problemas en:

1. Congeniar con otros.
2. Ser mandona
3. Escuchar a los demás (escuchar los de verdad, y no hacer como que escucho)
4. Matemáticas

La semana que viene quiero ocuparme en:

1. Congeniar con otros
2. Ser mandona
3. Escuchar de verdad a los demás
4. Estudiar matemáticas

Querida Jeanne,

Estás reflexionando maravillosamente bien sobre tus objetivos. Me alegra que hayas alcanzado tu objetivo de matemáticas. ¡Debes de haber estudiado muchísimo! En cuanto a los otros objetivos, creo que si tratas de alcanzarlos de a poco, paso a paso, podrías obtener mejores resultados. ¡Hazme saber cómo es esa “expresión especial” que aparece en tu rostro cuando estás escuchando de verdad.

♥ Srta G.

1-16

Querida señorita G:

Lo que me hizo pensar en lo de congeniar con otros es que soy mandona. Y sí realmente escucho a los demás, podré aprender muchas cosas. Y usted va a notar cuando estoy escuchando o no porque verá una expresión especial en mi cara. Me parece que no logré del todo la parte de congeniar, pero trate de no ser mandona ni pesada. Pero sí logré la parte de matemáticas, porque me saque 10 y sin ningún error en la prueba, así que, como premio mi mamá me preparó mi comida favorita isanguiches y camarones!

Jeanne

Querida Jeanne,





Estás reflexionando maravillosamente bien sobre tus objetivos. Me alegra que hayas alcanzado tu objetivo de matemáticas. ¡Debes de haber estudiado muchísimo! En cuanto a los otros objetivos, creo que si tratas de alcanzarlos de a poco, paso a paso, podrías obtener mejores resultados. ¡Hazme saber cómo es esa “expresión especial” que aparece en tu rostro cuando estás escuchando de verdad!

♥ Srta G.

CÓMO ENRIQUECER EL APRENDIZAJE Y HACERLO MÁS SIGNIFICATIVO

El portfolio del docente como responsable del desarrollo curricular y la evaluación es un documento valioso que muestra lo que han pensado y trabajado los docentes para hacer que el aprendizaje sea más rico y significativo para sus alumnos. Este portfolio revela la esencia de la práctica de los docentes y les permite a las personas ajenas al aula entender mejor las cualidades profesionales que hacen posibles la enseñanza y el aprendizaje. Julie COP, maestra de segundo grado del distrito escolar de Penfield, Nueva York, expone con elocuencia las ventajas de llevar un portfolio del docente como responsable del desarrollo curricular y la evaluación:

El portfolio de este año me sirvió para varias cosas:

-  *Me ayudó a determinar los resultados inherentes a crear una comunidad de lectoescritura y actualizar mi currículo. Mi programa tiene ahora mayor cohesión, pues cada componente se vincula a otro de un modo bien definido.*
-  *El proceso de crear una unidad de investigación me permitió ver la interrelación de todos los aspectos de una “buena unidad”. Los resultados, los indicadores, las r tablas de especificación y las evaluaciones contextualizadas funcionan en forma conjunta y global.*
-  *Constaté la importancia de coleccionar trabajos de los alumnos para crear puntos de referencia y modelos para ellos. Puede usar los trabajos recolectados el año pasado para mostrarles a los alumnos de este año qué es lo que considero un trabajo de buena calidad.*
-  *Lo más importante es que llevar un portfolio y participar en este proyecto me ha proporcionado el lenguaje y el marco de referencia necesarios para comprender mi práctica de la enseñanza en un nivel que no era capaz de definir antes de esta experiencia. Mi práctica ha contenido elementos q de buen diseño curricular desde hace años. Leo mucho, asisto a congresos y estoy orgullosa de mi capacidad de aprender. Sin embargo, antes tenía una sensación intuitiva de estar haciendo las cosas bien, pero no podría haber definido en público mi posición, como ahora estoy en condiciones de hacerlo. Aunque aún me queda mucho por aprender, tengo una clara noción de lo que es mi práctica profesional y de las razones por las que mi modo de enseñar es acertado.*

SCHLEMENSON, Aldo; “Antecedentes del Assessment Center y del abordaje a las organizaciones” Capítulo 6. Puesta en práctica. Desarrollo de un proyecto de evaluación del potencial” En: la estrategia del talento. México, Ed. Paidós. pp. 183-220.

6. PUESTA EN PRÁCTICA, DESARROLLO DE UN PROYECTO DE EVALUACIÓN DEL POTENCIAL

Los aspectos que atañen a la organización y a los recursos humanos fueron desarrollados a través de la práctica de la consultoría, en una diversidad de proyectos en organizaciones públicas y privadas: empresas industriales, bancos, empresas de telefonía, etcétera. En este capítulo haremos referencia a algunos de esos proyectos.¹

La experiencia nos señala que un proyecto de evaluación del potencial constituye una muy buena oportunidad para comenzar con un programa integral de cambio organización el. Como veremos a continuación, la propuesta supone un procedimiento claro, práctico y económico, fácilmente comprensible y asimilable por todos los miembros de la organización. En la forma en que aquí lo concebimos, ofrece una buena puerta de entrada a una variedad muy amplia de problemas en los que está involucrada la gente, el balance de los recursos existentes, su aprovechamiento, así como muchos temas fundamentales de estructura. En síntesis, permite arribar a un diagnóstico confiable para luego encarar acciones y desarrollar políticas y procesos de cambio den el ámbito de la organización y los recursos humanos.

En los organismos que forman parte de la administración pública, y en las grandes organizaciones en general, la consideración individualizada de la gente y la apreciación del mérito y del aporte individual constituyen un tema postergado pero a su vez prioritario para poder enfrentar la tendencia a la alienación y al anonimato característica de esos sistemas, que se convierten en burocráticos y formalistas.

Proponemos encauzar la evaluación sistemática de las personas que trabajan a través de dos sistemas que son pilares de la estrategia y la estructura de recursos humanos:

- a) *El sistema de evaluación de la efectividad, también denominado evaluación del rendimiento; y*
- b) *El sistema de la evaluación del potencial*

Implementar ambos procedimientos de evaluación contribuye a fortalecer el sistema gerencial, dado que son parte de las prácticas que debe realizar todo gerente para ejercer un liderazgo que lo posicione como verdadera figura de autoridad.

Cuando faltan ambos sistemas, o cuando no se implementan eficazmente, hecho que forma parte de una debilidad muy difundida de nuestros sistemas gerenciales o ejecutivos, se ve socavada su eficiencia. Esto, por lo general, está relacionado a su vez con la ausencia de la responsabilidad sistemáticamente implementada, tema ya considerado anteriormente. Cuando se carece de tales nociones, precisiones e implementaciones, que son parte de una estrategia de organización y recursos humanos, la autoridad del jefe o gerente no llega a

¹ Agradezco especialmente a la Lic. Beatriz Fontau, a la Lic. Claudia Fernández Vita, al Lic. Harald Solaas, al Lic. Norberto Di Paolo y al Lic. Alejandro Mazeira, por haber colaborado muy eficazmente en el desarrollo y en las conclusiones de este capítulo.

consolidarse plenamente. Tales carencias limitan el compromiso y diluyen la importancia de las decisiones gerenciales en distintos aspectos que hacen la conducción. Por lo tanto, consideramos oportuno recomendar la formulación y vigencia efectiva de los siguientes principios y prácticas. Nos referimos, básicamente, a la responsabilidad por:

- los resultados de los empleados a cargo;
- la evaluación equitativa y sistemática de la efectividad (comúnmente denominada rendimiento) y el potencial de esos empleados;
- la conducta laboral de los empleados.

Estos aspectos integran la definición de gerente en el capítulo sobre la estructura organizativa. Como puede advertirse, las responsabilidades acotadas están interrelacionadas y se potencian recíprocamente cuando la totalidad de las prácticas gerenciales mencionadas se encuentran en funcionamiento.

A veces ocurre que algunas de estas prácticas existen en un número significativo de organizaciones u organismos públicos, pero se verifica la falta de efectividad en su definición, en su implementación o en ambos aspectos. O, lo que nos resulta más importante aún, no forman parte de una política y de un marco referencial y operativo coherente.

Por eso, en los capítulos precedentes nos ocupamos de desplegar un enfoque de la organización y de la estructura integrando, que configure un verdadero sistema conceptual y operativo, que es el que le da sentido a las prácticas.

Es habitual que los sistemas puntualizados de evaluación se encuentren desprestigiados porque no se les implementa a pleno. Frecuentemente los gerentes no sumen esa tarea con el verdadero compromiso que requiere, y esto se refleja en la tendencia a calificar a todos por igual o a colocar a todos en los niveles más altos de las escalas elegidas. De esta forma, los sistemas no cumplen con la efectividad buscada. La ausencia de sistemas personalizados que comprometan a la discrecionalidad de los gerentes para realizar evaluaciones ecuanímes hace que la organización sea más impersonal y anónima, y así se vulnera la motivación de los empleados y se socava el sentimiento de trata equitativo que es posible desarrollar.

Existe también el supuesto de que la conducta de los gerentes carece de objetividad, ya sea porque no saben evaluar o porque sus decisiones están sujetas al *amiguismo*, las *trenzas* o los *compromisos*. También ocurre que se opta por la complacencia debido a que, por ejemplo, las remuneraciones en general son bajas. Muchas veces se otorgan posiciones de jefatura o gerenciales para compensar la imposibilidad de reconocimiento a través de la remuneración, dado que ésta se encuentra congelada. Entendemos que todos estos casos y situaciones constituyen síntomas o efectos de la ausencia de una verdadera *política ecuaníme, coherente, confiable, que restablezca la equidad, la justicia y que reconozcas como valor fundamental el mérito individual*.

Resumiendo, la carencia puntualizada favorece la impersonalidad de los sistemas vigentes; se suma a esto la falta de un liderazgo gerencial proactivo sólidamente instalado. A la postre, los sistemas de administración de recursos humanos rehacen *menos confiables y transparentes* para los empleados y para el público en general y así se desprestigian.

Una concepción que favorezca la transparencia y la confiabilidad de los sistemas gerenciales, que propugne la efectividad del rol gerencial, el trabajo en equipo, un buen servicio a los usuarios o clientes y que se proponga movilizar el grupo humano hacia el cumplimiento de resultados específicos, requiere, entre otros cambios muy importantes, que

los gerentes se comprometan en la forma de decisiones en relación con la evaluación de la gente. Ésta es la forma, a su vez, de erradicar un aspecto de la corrupción a través de la línea ejecutiva. Cuando el jefe o gerente es responsable por la conducta laboral de la gente, los envíos de las conductas individuales pueden ser controlados a través del sistema gerencial, que cumple con una finalidad preventiva.

No hay que dejar libradas las buenas prácticas a las actitudes de los buenos gerentes. Con pretendemos demostrar en esta obra, las cualidades de una buena organización se refuerzan cuando todas las conductas están respaldadas por la organización y por la incorporación sólida del concepto de responsabilidad bien instaurado, cuando se legitiman las prácticas de evaluación como un requisito inalienable que todo gerente debe realizar en forma concreta. Para ello, los gerentes deberán conocer y centrarse en aquellos aspectos esenciales de la conducta humana determinantes del rendimiento y del nivel potencial más alto en el que es posible resolver problemas. El mensaje institucional claro y con sentido unívoco ayudará a la implementación de las buenas cualidades de una organización.

En síntesis, todo intento de mejora y búsqueda progresiva de eficiencia de la organización debe compensar las tendencias alienantes de los modelos burocráticos e instaurar el principio de la responsabilidad gerencial, que involucra una variedad de prácticas que el gerente debe realizar y que aquí describimos con algún detalle. Entre otras, la responsabilidad por la evaluación.

LA FUNCIÓN DE LOS RECURSOS HUMANOS EN EL FORTALECIMIENTO DE LOS SISTEMAS EDUCATIVOS

En los grandes organismos, los niveles superiores a veces tienen un conocimiento limitado de un reducido número de empleados con los que les toca interactuar. La posibilidad de que un gerente, ubicado en el estrato IV o V, pueda considerar para una posición de conducción a los empleados ubicados, por ejemplo, en estratos II o III, se ve restringida por la falta de conocimiento personalizado de la gente.

De esta forma, los sistemas ejecutivos de conducción se vuelen aún más impersonales, en especial para los profesionales y empleados que, teniendo los talentos necesarios, se ubican en los estratos inferiores en unidades organizativas descentralizadas. Como resulta obvio, este factor es desalentador, y va a contrapelo de las modernas teorías de la motivación.

Hemos verificado que jóvenes profesionales reclutados en las universidades tras exigentes exámenes de ingreso pueden tardar, en un organismo público o en una empresa grande, un tiempo excesivo para ocupar una posición de jefatura o gerencial. Este tiempo no guarda relación con el que insume la maduración y el desarrollo individual, por cierto muchos más breve. De hecho los jóvenes profesionales con un buen potencial de desarrollo están en condiciones de ocupar una jefatura o gerencia de estrato II o III, o aún más alta en un lapso breve. Muchas empresas privadas que aspiran a capitalizar el talento de sus empleados se esfuerzan por desarrollar sistemas operados por las gerencias medias o superiores. Hemos participando en decenas de programas de jóvenes profesionales que en épocas de desarrollo de la economía buscaban evaluar el alto potencial para consolidar un *pool* de talentos que permitiera satisfacer las necesidades de la estructura y del desarrollo. Los jóvenes profesionales que ingresan a través de programas especialmente diseñados pueden, en un término de dos años como máximo, estar en condiciones de desarrollarse

con autonomía y hacerse cargo de proyectos especiales en posiciones gerenciales. En épocas normales, los que no están en condiciones de desarrollarse en el plazo mencionado son desalentados a continuar e impulsados a buscar otros destinos. Lo que acabamos de decir suena irreal en las condiciones actuales de gran recesión en la Argentina. Pero es posible observar que, en épocas de desarrollo como las que seguramente volveremos a retomar, en países en los que predomina el empleo abundante, los jóvenes profesionales de alto potencial tienden a desertar cuando su talento no es reconocido o cuando no ocupan una posición acorde con el mismo. Aun en épocas recesivas de desempleo y su empleo, los más talentosos encuentran oportunidades más rápido. El fracaso de los programas para retener a los jóvenes de alto potencial obedece a la ausencia de instrumentos ágiles para ubicar con bastante rapidez a los candidatos en posiciones y proyectos según su capacidad. La falta de relaciones tutoriales con gerentes de los niveles superiores acentúa la tendencia.

Queremos decir que lo que se necesita hacer para alentar y retener a los jóvenes profesionales de alto potencial, que constituyen el reservorio de talentos de las grandes organizaciones, es ubicarlos rápidamente en proyectos que representen para ellos un interés y un desafío y que estén a cargo de gerentes de alto rango (por ejemplo de estrato IV), quienes les van a mostrar un plafón más alto para su desarrollo.

En la administración pública y también en muchas empresas probadas, la ausencia de sistemas de recursos humanos eficaces hace que el talento humano, que es el valor agregado más importante para la innovación, se vea desaprovechado. La estabilidad que ofrecen muchos organismos –valor muypreciado en épocas de recesión, desempleo e incertidumbre—favorece la permanencia, pero no obstante, la rigidez de los sistemas imperantes frustra a los que tienen mayores posibilidades de desarrollo por la falta de oportunidades para poner a prueba su potencial. Las organizaciones se ven privadas así de l talento, el entusiasmo y la idiosincrasia proactiva de una población potencialmente capaz de movilizar la cultura. De esta forma, también se desalientan la ejecutividad y la flexibilidad para enfrentar los cambios, cualidades que permiten adecuarse a las exigencias de la época.

Cuando los sistemas de recursos humanos están afectados por la rigidez no garantizan que las posiciones de mayor responsabilidad estén ocupadas por los individuos más talentosos. Las limitaciones, en términos de capacidad individual de las principales figuras de conducción, se constituyen en un obstáculo fundamental para el desarrollo organizaciones. La totalidad del sistema conformado por individuos se niveles abajo si las posiciones ejecutivas máximas no están ocupadas por quienes cuentan con la capacidad adecuada.

La complejidad más grande que se afronta es la que puede manejar el jefe. *Por ello, el valor agregado fundamental es el nivel de capacidad del ejecutivo principal de una unidad organizativa. De su talento surgen los recursos creativos para resolver situaciones de crisis.*

Como lo señala el *principio de Peter*, muchos directores que ocupan posiciones estratégicas y tienen responsabilidades en unidades organizativas de hasta IV o V estratos han llegado a su nivel de incompetencia. La totalidad del sistema sufre esta limitación nivelándose hacia abajo.

Una reglamentación general y común para todos pretende limitar los efectos negativos de la discrecionalidad. Para ello, en el pasado se recurrió a cláusulas de promoción o

ascenso automáticos que determinaban las promociones por antigüedad que en muchos sistemas administrativos reemplaza al mérito. Esto favorece una modalidad retardataria y conservadora que frena el desarrollo estratégico de los organismos e impide que se nutran de talento, que no se correlaciona con la edad.

No sólo la rigidez o falencia de los sistemas de recursos humanos, sino también la rigidez del contrato laboral, limitan la discrecionalidad de los jefes y gerentes para tomar decisiones y medidas que atañen a sus empleados a partir de la realización de las evaluaciones ecuánimes. En la administración pública, las transgresiones a la ética por parte de algún empleado, por ejemplo, termina siendo protegidas por un sistema impersonal, que exige una serie complicada de sumarios y procedimientos judiciales para decidir sanciones frente a algún desvío grosero, lo que desalienta a cualquier gerente para tomar decisiones preactivas.

En función de estas consideraciones, lo que se desea es imprimir un cambio significativo para lograr una mayor eficiencia, un mejor servicio al cliente o al público en general, la estrategia de desarrollo es comenzar por el principio, es decir por el talento. Esto significa valorizarlo en primera instancia dado que conforma el capital humano con que el que se cuenta; se trata de una consideración seria y sistemática de las necesidades y capacidades de la gente, principal factor de movilización de las grandes organizaciones. Una vez que reconocemos la importancia del talento, debemos detectarlo, realizar una búsqueda sistemática, y el paso siguiente es *liberarlo*, es decir, darles capacidad discrecional a los gerentes para tomar decisiones creativas y no previsibles, en relación con sus colaboradores inmediatos y en relación también con los problemas que tienen que resolver. Todo esto, para no reducirse al ejercicio del libre albedrío, necesitará estar encauzado y limitado por un sistema sólido de responsabilidades gerenciales que establezca como sello el deber de rendir cuenta por las acciones emprendidas y por los resultados obtenidos de los equipos a cargo.

Como ya señalamos, un desarrollo organizacional de esta naturaleza presupone un enfoque sistémico e integral, en el cual se encuentren balanceados la estrategia global, la estructura adoptada, los estratos organizacionales necesarios y la correcta consideración y ubicación de la gente.

Cuando la evaluación del potencial es realizada, por ejemplo, por expertos externos y no es integrada como sistema a un modelo coherente de organización y gerenciamiento operado por el propio cuerpo gerencial, la herramienta constituye un mero paliativo tendiente a suplir la presunta inhabilidad de los gerentes para realizar evaluaciones confiables. La presunción se convierte en una verdadera profecía que se confirma en la práctica, a través de decisiones centralizadas que limitan la discrecionalidad de los gerentes, quienes terminan por lavarse las manos y eludir la responsabilidad por la evaluación, que debería ser indelegable.

Si bien la perspectiva de comprometer a los gerentes y de aumentar su discrecionalidad favorece la transparencia organizacional, y los proyectos de *evaluación del potencial* y de *la evaluación de la efectividad personal* son fundamentales, estos dispositivos no pueden ser considerados como un remedio de sí mismos. Debe integrárselos al marco de referencia adoptado, a la multiplicidad de aspectos que hacen a una transformación global y profunda de la organización.

Desde el punto de vista táctico, el proyecto de *evaluación del potencial* constituye una buena oportunidad para encarar un tema acuciante de recursos humanos, aunque no es el

único. Es una forma de adelantar camino en pos de la transparencia en el área que atañe a la gente.

Por su parte, el sistema de *evaluación de la efectividad* constituye una herramienta para alentar e incentivar a la gente, instituyendo el principio del mérito y reconociendo en función del rendimiento, y es un recurso que fortalece la propia responsabilidad del gerente. Si bien este sistema de evaluación del potencial, la evaluación de puestos y el sistema de remuneraciones (que parte de una escala salarial equitativa). Este enfoque exige que, inicialmente, las definiciones estratégicas sean formuladas por la conducción superior y aprobadas por el ejecutivo principal. En resumen, la forma que desde la perspectiva ofrecemos para comenzar a ocuparnos de la gente consiste en dos sistemas que permiten verificar la correspondencia entre el nivel del cargo que se ocupa y la capacidad de cada uno de los evaluados. Estos sistemas son la evaluación de la efectividad personal y la evaluación de la capacidad potencial. Ambos representan los pasos preliminares para promover y ayudar en su desarrollo a aquellos que están en condiciones de ocupar posiciones superiores y que, muchas veces, fueron dejados de lado.

LA EVALUACIÓN DE LA CAPACIDAD REALIZADA POR GERENTES

En este punto comenzamos a adentrarnos en la práctica de la evaluación y, en especial, aquella realizada por los gerentes del sistema ejecutivo. Es decir, la que se practica a través de la línea ejecutiva. En este caso, las evaluaciones son realizadas por los gerentes, y por los gerentes de los gerentes, y atañen a colaboradores de los colaboradores directos de los segundos. Las definiciones de capacidad desarrolladas en el capítulo anterior presentan la ventaja de haber sido formuladas dentro del contexto de una organización ejecutiva que involucra el trabajo de gerentes.

Como ya lo señalamos en el capítulo sobre la estructura organizativa, el sistema de trabajo gerencial supone un eslabonamiento mínimo de tres niveles:

1. el gerente del gerente;
2. el gerente directo que depende del anterior;
3. el rol que está siendo evaluado (colaborador o subordinado inmediato del gerente anterior).

Algunas premisas para realizar evaluaciones confiables

Para que los sistemas gerenciales funcionen en forma correcta deben garantizarse las siguientes condiciones:

- El gerente debe estar trabajando en un nivel de abstracción superior (correspondiente al estrato siguiente), diferente e identificable respecto de su colaborador inmediato.
- A los fines de establecer parámetros objetivos para la evaluación de colaboraciones directos e indirectos, es indispensable referirse a las tareas y a la estructura de roles dentro de los que se desarrolla el proceso de trabajo. Los roles se distribuyen en estratos, y éstos representan un nivel de complejidad y abstracción determinado.
- Hay una lógica implícita en este planteo: resulta sencillo y ventajoso describir en forma objetiva los distintos niveles posibles de trabajo, y luego referir a ellos los

distintos niveles de complejidad de la capacidad potencial. De aquí surge también que existe una forma de evaluar la capacidad potencial, y es refiriéndola a los distintos niveles de complejidad de las tareas que se realizan en una organización.

- Es importante destacar que en los sistemas ejecutivos gerenciales, los gerentes y los gerentes de los gerentes ejercen una apreciación intuitiva acerca de la complejidad de las tareas, y sin capaces de relacionarlas con la complejidad de los procesos mentales que los colaboradores directos y los colaboradores de sus colaboradores utilizan para resolverlas.

O sea que no les resulta difícil referirse a la complejidad de los procesos mentales y, de hecho, responden con mucha precisión

Cuando se les convoca a evaluar el potencial si está bien establecido el marco de referencia de los estratos organizacionales. Esto también requiere que se les explique con mucha claridad cada uno de los términos de referencia implícitos en las definiciones, lo que compromete un esfuerzo didáctico significativo por parte del experto, previo a la introducción de un sistema de evaluación de la capacidad potencial actual y futura.

Es menester no sólo explicar cada una de las definiciones y de los componentes, sino también explicitar las diferencias entre los distintos conceptos involucrados. Cuando esto es captado y los gerentes aprehenden el sentido unívoco de las definiciones y asimilan el lenguaje de los estratos organizacionales, pueden abocarse con facilidad e interés a la realización de la tarea evaluativa propuesta. Por cierto que esta actividad ha demostrado ser sumamente interesante para los propios gerentes.

- ◆ Debemos recordar que evaluar es emitir un juicio gerencial, en última instancia, subjetivo. Es tomar una decisión acerca de la ponderación de un atributo de la conducta de colaboradores directos o indirectos (es decir, de personas ubicadas a uno o dos estratos de distancia del propio). ◆

Los gerentes, cuando asignan tareas y cuando establecen un intercambio para el ajuste del desempeño, como parte de la conducción del proceso de trabajo de sus colaboradores, constantemente *están realizando y emitiendo juicios gerenciales* que involucran una evaluación del subordinado, lo que constituye un aspecto central de la toma de decisiones que cada gerente realiza. Es necesario admitir que, en el acto de conducir, constantemente se toman decisiones que tienen que ver con la evaluación de la conducta de los colaboradores directos.

La posibilidad de referenciar estos juicios a mejores inequívocos de la realidad externa – los estratos y los niveles de complejidad-, y luego ajustar los juicios mediante la confrontación de las subjetividades es un procedimiento válido que facilita mucho las evaluaciones. Dicho de otra manera, cuando el gerente comprende con claridad las definiciones, es decir, qué es lo que se está evaluando, y cuando comprende el trabajo que el colaborador realiza, en términos de la posición del rol al que pertenece una jerarquía de complejidad, la tarea se le presenta como sumamente clara y precisa. Es allí cuando puede demostrar su conocimiento intuitivo acerca de las personas. Este conocimiento viene de la práctica y se fortalece en la asignación de tareas. Una vez más, el hecho de ser responsable por el desempeño de un colaborador directo liga al gerente de una manera muy especial con el comportamiento y con los resultados de su colaborador. Se crea una doble

dependencia que los une en la relación de trabajo y que tiene como telón de fondo, a su vez, la apreciación del gerente del gerente.

Como lo venimos señalando, el horizonte temporal del rol o ITD ofrece un dato para la observación objetiva y para la determinación del nivel. Este factor tiempo, que es susceptible de ser descubierto a través de una indagación referida a las tareas que un gerente asigna a un subordinado, nos ofrece una medida objetiva del peso específico del rol en cuestión.²

Esta descripción sugiere que el *nivel de capacidad de trabajo puede ser expresado en términos de los lapsos temporales que un individuo maneja*. Si esto es así, la complejidad de los procesos mentales que posee puede ser medida en términos del tiempo hacia el futuro que puede manejar, organizando proyectos y actuando en ellos con coherencia para producir algo.

Tenemos, entonces, que el horizonte temporal de un individuo está dado por aquella meta que tiene el plazo de consumación más lejano. Se deduce de la tarea planeada de más largo plazo y forma parte de su presente activo; es el punto más lejano hacia donde mira en un momento dado. La distancia depende de su nivel de capacidad y, por lo tanto, es una manifestación objetiva de esta capacidad.

Por medio de la aplicación sistemática del horizonte temporal y del desarrollo conceptual de una jerarquía de niveles de complejidad, se pueden establecer límites nítidos que representan puntos de corte, en términos de la tarea extendida más larga, se produce un cambio cualitativo en la naturaleza de la misma y de su complejidad. Como lo hemos señalado, los puntos de corte se presentan al término de un día, tres meses, un año, dos años, cinco años, diez años, veinte años. El pasaje de un límite representa la entrada a un nuevo estrato.

◆A través de la asignación de tareas, el gerente delega sólo aquellas que el empleado puede cumplir, y esto lo hace en función de la apreciación de su capacidad, que en forma implícita o explícita realiza ◆

El lapso de discrecionalidad máximo de una persona y la observación del grado de cumplimiento satisfactorio de las asignaciones correspondientes a ese lapso, ofrecen la posibilidad de evaluar en forma objetiva y efectiva la capacidad de un individuo en su trabajo. El gerente es intuitivamente consciente de esos lapsos de autonomía temporal que asigna. Puesto que, de otra forma, no podría hacerse responsable por los resultados de sus empleados. Cuando el gerente asigna trabajo a un empleado, en el mismo acto de la delegación, implícita o explícitamente, también emite un enunciado acerca de los límites prescriptos, los métodos que deben ser empleados y la calidad esperada del trabajo, junto con la especificación del plazo máximo de realización.

Sin estos elementos, el trabajo realizado por el empleado se contraría fuera de contexto y desvinculado de la realidad.

² Para la medición del time-span del rol en una organización ejecutivo-jerárquica, veáse el capítulo 3, “La estructura organizativa”.

PROCEDIMIENTOS PARA ESTABLECER LOS SISTEMAS DE EVALUACIÓN

Determinación de los estratos requeridos

Como ya se ha dicho, un punto clave de este enfoque es tener muy en claro los estratos organizacionales requeridos en forma previa a las evaluaciones. Para ello se necesita:

- a) Determinar los niveles ejecutivos requeridos o estratos. Determinar el nivel de los roles involucrados en la evaluación, utilizando el instrumento del ITD y el esquema antes expuesto de los niveles de complejidad necesarios que surgen del modelo estratos.
- b) Detectar y describir la organización existente y, dentro de ella, la superposición de roles, las situaciones de comprensión (roles supuestamente subordinados que se ubican dentro de un mismo estrato); la excesiva distancia entre roles que supuestamente mantienen relación de dependencia directa (cuándo éstos se ubican a más de un estrato de distancia entre sí); las líneas de dependencia o autoridad poco claras.

Los casos de comprensión revelan una falta de distancia óptima desde el punto de vista del diseño, que impide que se cumpla el liderazgo efectivo. Cuando se define un organigrama, los roles de gerentes deben estar en el estrato inmediato superior al de los colaboradores inmediatos.

Todas estas situaciones son consideradas síntomas o anomalías institucionales que obstaculizan el funcionamiento del sistema ejecutivo y dificultan la aplicación del ejercicio de evaluación del potencial. Imprecisiones como las puntualizadas impiden que los juicios evaluativos que se formulan en el proceso de evaluación del potencial tengan referentes claros y unívocos. Para que las líneas de autoridad y la distancia óptima estén bien diseñadas desde el punto de vista de la relación jerárquica entre los roles, la organización requerida debe coincidir con la existente. Por ello, sería del corregirlas en forma previa o, ni no, tenerlas en cuenta para ajustar el procedimiento.

Evaluación de la efectividad personal

Utilizamos la expresión *efectividad personal* en lugar de *rendimiento* para referirnos al grado de cumplimiento y de compromiso que los empleados demuestran al aplicar su esfuerzo y su capacidad al logro de las tareas y objetivos asignados. Para trabajar deberán enfrentar las dificultades que se les presentan haciendo uso de los recursos disponibles que actuarán como limitantes del flujo de conducta desplegado.

El que evalúa la efectividad del ocupante del rol es el gerente directo, quien tiene la responsabilidad de juzgar la efectividad de su colaborador para llevar a cabo las tareas y enfrentar los problemas y oportunidades que se le presentan al hacer su trabajo. Es decir, el gerente evalúa cuán correctamente el subordinado alcanzó el resultado.

Como ya se ha dicho, esta evaluación global compromete un juicio de valor gerencial. Emitirlo constituye una de las decisiones más importantes que el gerente directo debe tomar.

Para llegar a la conclusión acerca de la evaluación, el gerente no se basa en datos derivados de resultados cuantitativos ni de la consideración de factores o competencias particulares. Este punto es crucial, y con respecto a él discrepamos con muchos de los esquemas

vigentes. Los resultados cuantitativos no dependen de la efectividad aplicada ni del buen trabajo realizado por los empleados, sino que están afectados por los recursos disponibles, entre ellos los financieros y por las condiciones del mercado, y estos aspectos no pueden ser modificados por los empleados.

Las circunstancias bajo las cuales éstos trabajan son relativamente imprescindibles, y por lo tanto la discrecionalidad individual no las afecta en forma directa cuando se trata de cumplir con una asignación determinada.

¿como podría juzgarse por los resultados a un vendedor en el contexto de un mercado recesivo o una crisis económica como la que vivimos, siendo que los principales factores que hacen a la venta o a la recaudación escapan a su manejo? la verdadera capacidad aplicada al trabajo debe ser evaluada en forma independiente de los resultados, dado que se pueden obtener altos resultados con una baja efectividad aplicada, al revés, en situaciones de mucha dificultad se pueden trabajar muy bien enfrentando problemas difíciles con un bajo resultado promedio.

La evaluación compromete un juicio global. Consideramos artificial e innecesario recurrir a un sistema de factores para realizar las evaluaciones. La evaluación de la capacidad aplicada actual es una forma de evaluar la conducta de la persona total, inserta en un puesto y en un sistema laboral determinado esta ponderación, por ende, reclama un juicio holístico.

Objetivos de la evaluación personal

A continuación señalamos los principales objetivos que pretendemos alcanzar con la introducción del sistema de evaluación de la efectividad.

- Contar con una evaluación personal e individualizada, que refleja el rendimiento de cada empleado en un período de tiempo, como mínimo anual.
- Permitir que los empleados tengan un *feedback* o devolución de su gerente directo, que explicita cómo es evaluada (juzgada) su efectividad personal en el trabajo. Esta evaluación es un recurso para comenzar un proceso de diálogo que lleva implícito el señalamiento de áreas de mejora o puntos débiles, y el reconocimiento del esfuerzo y la capacidad aplicada en el trabajo.
- Dar la oportunidad a los empleados o colaboradores directos de un gerente para expresar sus puntos de vista y discutir el resultado de la apreciación de su efectividad personal.
- Comenzar a brindar la oportunidad de realizar un *coaching* sistemático, que constituye una forma de entrenamiento en el rol de trabajo.

Los procesos como la evaluación de la efectividad personal son parte de una práctica cotidiana que se integra con fluidez y naturalidad a la relación diaria de un empleado con su gerente. La práctica, en la medida en que es incorporada en la organización, se transforma en una destreza, que por definición se realiza en forma natural. Es parte de un diálogo regular centrado en la realización del trabajo que surge de las asignaciones recibidas, y que fortalece el vínculo de un colaborador con su gerente. No está desconectada del resto de las prácticas que el gerente debe desplegar en relación con sus colaboradores directos. En efecto, está en particular muy ligada con la asignación de tareas; sino están claros los objetivos de la tarea y el perfil del rol que incluye las responsabilidades por las cuales el

empleado deberá rendir cuenta, mal se podrá evaluar la efectividad aplicada a los resultados esperados.

En el logro de las tareas asignadas, insistimos una vez más, los empleados son responsabilidades por el uso de sus habilidades, conocimientos, experiencia, compromiso e interés provenientes de la valoración de lo que se les encomienda realizar. Cuando las asignaciones no están claramente establecidas o comprendidas, el empleado podrá retomar el dialogo con su gerente y solicitarle todas las aclaraciones que necesite. Esta recurrencia inmediata y anticipada al gerente, atendiendo a las dificultades y problemas que se presentan, es un derecho y una obligación cuya práctica se debe estimular.

Es frecuente que la evaluación se transforme en una práctica que se realiza rutinariamente una vez por año, sin entusiasmo y mecánicamente. Es aceptada por los jefes como una imposición que viene de afuera, o respetada por provenir de una área colega, generalmente recursos humanos. Podemos decir que en estos casos el compromiso por parte de la línea ejecutiva es débil, y que el sistema no se asimiló a la cultura gerencial de la organización. Para que la práctica sea efectiva, en cambio, deberá integrarse a otras prácticas que en su conjunto sustituyan una estrategia de la de

Dirección y, por lo tanto, establecerse como parte de las políticas de la empresa. Deberá tratarse de una estrategia global, integrada a un cambio cultural también global. Cada vez que un empleado cumple con una asignación recibida, se presenta la oportunidad de iniciar un diálogo constructivo que lleva implícita una decisión evaluativo por parte del jefe.

Evaluación anual de la efectividad personal

La evaluación anual de la efectividad personal se transforma en un resumen de los intercambios mantenidos durante el año, los cuales habrán sido registrados en sus aspectos más relevantes por el jefe directo. La evaluación sintética anual es proporcionada por el gerente en el momento en que debe tomar una decisión que trae aparejado un ajuste anual de la compensación por méritos para el empleado, dentro de las pautas preestablecidas. Como parte de este proceso, los gerentes juzgan la capacidad aplicada de cada empleado de acuerdo con el rol en el que está trabajando. El esquema propuesto con Jaques ofrece seis escalones dentro de cada rol, y el gerente ubica a cada uno de sus empleados en ellos. Esta evaluación permite luego fijar y asignar un incremento establecido anual por méritos.

De hecho cuando las evaluaciones no se realizan de forma ecuaníme, cuando a través de cualquier subterfugio el gerente elude la responsabilidad de tomar estas decisiones evaluativos que apuntan al reconocimiento de las diferencias entre los empleados, la consecuencia es el desdibujamiento de su autoridad. En nuestra experiencia, esto no acontece cuando media un verdadero compromiso, o sea, cuando esta comprometida la línea, comenzando por el ejecutivo principal de la organización en cuestión. En este caso, el área de recursos humanos interviene por delegación expresa del ejecutivo principal, brinda un servicio de apoyo y se ocupa del monitoreo del sistema.

En síntesis, como dijimos en la introducción, la evaluación de la efectividad personal ejercida por el gerente conforma una de las decisiones más importantes que hacen el liderazgo gerencial efectivo; y la consecuencia de una aplicación correcta, se aprecia en el fortalecimiento de la autoridad del gerente.

Participación del gerente del gerente

El gerente del gerente esta comprometido y participa en el proceso de evaluación de dos maneras: a) cuando el empleado ejerce su derecho de apelación ante la evaluación que no considera justa; y b) en el proceso de equilibración.

- a) *Apelación de la evaluación.* Si alguien considera que no ha sido tratado en forma justa por su propio gerente, puede apelar al gerente del gerente, y si persiste su insatisfacción, puede acudir a niveles más altos en la organización. Este mecanismo se práctica con el conocimiento del gerente inmediato y de todos los niveles gerenciados involucrados. El área de recursos humanos deberá estar notificada de la existencia de una apelación solicitada por un empleado.
- b) *Equilibración.* El gerente del gerente tiene la responsabilidad de realizar una reunión para equilibrar los juicios de efectividad personal realizados por sus gerentes subordinados directos, es decir, para realizar comparaciones colaterales. De esta forma, asegurará la existencia de un trato equitativo y consistente para todos los empleados de su área.

Procedimiento para la evaluación de la efectividad personal

Como consecuencia de estos procesos, los gerentes van conformando sus propios juicios acerca de la efectividad con la que trabajan sus empleados. El *feedback* continuo les aporta datos acerca de su desempeño teniendo en cuenta las circunstancias, o recomendaciones críticas según corresponda, esto forma parte del mecanismo de *coaching* que práctica el gerente con todos sus colaboradores inmediatos.

Es de buena práctica que el gerente reúna durante el año incidentes críticos significativos que surgen de las revisiones periódicas de la efectividad personal. Estas notas tendrán particular relevancia al finalizar su período completo de evaluación, y de esta forma la instancia de la devolución no ofrece sorpresas para el empleado. Si bien se recomienda realizar una evaluación formal por año, la frecuencia puede ser mayor y puede haber intercambios durante el año que no impliquen una documentación formal.

Pasos de la evaluación

1. *Los gerentes juzgan la capacidad aplicada de cada colaborador inmediato.* Una vez por año. Los jefes realizan una evaluación formal de cada uno de sus empleados inmediatos, teniendo en cuenta su capacidad aplicada acorde con los componentes mencionados. En esta evaluación, que constituye un juicio de valor, es ubicada dentro de una escala de seis niveles (Jaques, 2000).

Resulta de buena práctica que los gerentes estimulen la autoevaluación de sus empleados, que podrían expresar sus preocupaciones referidas a su rol. Los puntos más significativos de este intercambio serán consignados en un informe.

El gerente juzgará si sus empleados están trabajando en forma acorde con las expectativas del rol actual, si lo están haciendo en un nivel superior o inferior.

2. *Rating de la efectividad personal.* Se intentará evaluar la capacidad de los empleados estableciendo una comparación entre todos aquellos que forman parte de un mismo plantel.

A fin de establecer las evaluaciones el gerente podrá formularse las siguientes preguntas: *¿el empleado X está desplegando un nivel de capacidad consistente con la capacidad requerida por el rol...? ¿Superior a la requerida por el rol...?*

Si el empleado está ejerciendo un nivel de capacidad consistente con los requerimientos del rol, ¿lo está haciendo como alguien que puede ser ubicado en la mitad superior del rol, o como alguien que puede ser ubicado en la mitad inferior del rol? En caso de encontrarse en la mitad superior, ¿se encuentra por encima de esta mitad, por debajo o en el centro?

El *rating* por mérito para cada empleado se ingresa en el formulario del gerente y se entrega al gerente del gerente para su equilibración. Previamente se habrán trabajado los casos con los empleados cuya capacidad aplicada actual se encuentra por debajo del nivel requerido. El gerente del gerente y el área de Recursos Humanos están al tanto de estas situaciones.

3. *Equilibración.* El gerente del gerente resume todos los *ratings* de evaluación de los empleados de sus propios gerentes subordinados y los ingresa en una misma planilla de equilibración. La esencia de la reunión de equilibración consiste en advertir y considerar situaciones en las que el gerente ha sido excesivamente indulgente o generoso o excesivamente estricto.

El gerente del gerente mantiene una reunión de equilibración con todos los gerentes inmediatos que dependen de él para asegurarse de que las evaluaciones de efectividad han sido realizadas en forma justa y honesta. Compara los *ratings* utilizados por cada jefe para asegurar que haya una unidad de criterio entre todos los involucrados.

4. *Reunión de evaluación anual.* Una vez realizada una discusión detallada sobre la evaluación de la efectividad, se otorga a cada empleado la oportunidad de revisar el documento de evaluación personal para que incluya sus comentarios sobre el contenido. Luego el empleado revisará el documento y lo firmará una vez terminada la reunión puede documentar sus razones por escrito y solicitar otras discusiones con el gerente inmediato con el gerente del gerente. El gerente podrá hacer modificaciones como consecuencia de estos intercambios adicionales.
5. *Reconocimiento por mérito.* Las decisiones del jefe sobre el reconocimiento anual por mérito deben estar sustentadas por la evaluación de la efectividad personal. Las sumas aprobadas de reconocimiento se comunican al empleado en una reunión individual con una fecha fijada en el transcurso del año.

EVALUACIÓN DE LA CAPACIDAD POTENCIAL ACTUAL

Como se expuso en el capítulo anterior, se entiende por capacidad potencial actual, que incluye los conocimientos, habilidades y destrezas así como la valoración que el empleado hace del trabajo que realiza, la evaluación del potencial actual abstrae la esencia de la contribución individual (la complejidad de los procesos mentales), dejando de lado los otros elementos antes enunciados y que son susceptibles de ser adquiridos a través de la experiencia o el entrenamiento. Evaluar el potencial actual es una tarea de máxima

importancia que realizan los gerentes en todos los niveles, ya que constituye una de las herramientas más fuertes para el futuro de la organización

Para pasar al plano de la implementación práctica de un sistema de evaluación del potencial podemos consignar, por una parte, el propósito global y, luego, los objetivos específicos.

El propósito global que inspira este proyecto es:

- ♣ favorece el uso pleno del potencial individual;
- ♣ hacer regir el principio del mérito en la organización, valorizando el talento para cubrir posiciones estratégicas ;
- ♣ promover y desarrollar a todos los empleados teniendo en cuenta la capacidad potencial dentro de las oportunidades existentes;
- ♣ lograr una visión global y un balance del *pool* de talentos de la organización.

Los objetivos específicos son:

- ♣ evaluar la capacidad potencial actual de aquellos que ocupan posiciones, comenzando por los de mayor nivel jerárquico y cubriendo la totalidad de los empleados mediante un sistema que abarca estrato organizacional;
- ♣ realizar una apreciación global, estrato por estrato, de la capacidad potencial; hacer un balance de los recursos humanos disponibles en función de los niveles de los puestos y de las personas involucradas en ellos;
- ♣ detectar un grupo de individuos con talento y potencial para ocupar puestos en el momento actual y para planificar la cobertura futura de puestos en función de las expectativas de desarrollo de la organización.

Las dos estrategias para la evaluación del potencial

En términos generales, pueden plantearse dos posibles estrategias para encarar el proceso de evaluación del potencial, que fueron contempladas en el capítulo anterior:

- a) evaluación individual del potencial hecha por consultores externos a través del análisis del discurso. La metodología que ha sido desarrollada por Jaques y Cason (1994) permite detectar cuatro formas de procesamiento mental que se manifiestan en niveles progresivos de complejidad en el procesamiento de la información.
- b) Evaluación a través de la línea ejecutivo-jerárquica, en cascada, comenzando por el ejecutivo principal, evaluando a los subordinados de sus subordinados. Esta información luego es checada con el jefe directo.

En uno de nuestros proyectos recientes pensamos implementar una técnica mixta que pudiera representar un doble control de las evaluaciones realizadas. Se consideró, en un principio, que la falta de entrenamiento y hábito o la eventual subjetividad por parte de los jefes, podían afectar la ecuanimidad de las evaluaciones e invalidar el procedimiento. Contribuye a esta presunción el supuesto de que la convivencia prolongada entre gerentes y colaboradores dentro de una organización hace que se generen vínculos primarios, alianzas implícitas espontáneas e inconscientes, y de que éstas limitan la aproximación objetiva y

racional a las decisiones evaluativas. Si bien esto tiende a ocurrir, el fenómeno no depende de un hecho espontáneo e independiente.

Condiciones para el funcionamiento efectivo del sistema

De acuerdo con nuestra experiencia, que se repite cada vez que implementamos un sistema de esta naturaleza, los *jefes o gerentes en general son ecuanimes cuando:*

- ♣ Las políticas y los sistemas que se les proponen son claros, sensatos y alentados por un principio de equidad y de justicia;
- ♣ Los sistemas están oficialmente sancionados;
- ♣ La aplicación fluye armónicamente a lo largo de la línea;
- ♣ El ejecutivo principal comienza a aplicarlos con seriedad y dedicación, comunicando con su propia práctica que la consideración individualizada de la capacidad de la gente constituye un valor de la organización alentado por el sistema gerencial.

Dadas estas condiciones, el ejemplo cunde. La evaluación sistemática considerada como una práctica sancionada se transforma en un mensaje alentador para realizar una evaluación ecuanime y sensata.

Cuando las figuras ejecutivas centrales adoptan el sistema, y lo realizan seriamente ellas mismas, la aplicación fluye sin problemas y se transforma en un método que da lugar a coincidencias significativas entre aquellos que participan. Cuando esto no se produce, esta prevista la búsqueda del consenso a través e la negociación.

El propio Jaques es muy enfático en sus indicaciones acerca de cómo usar los procedimientos y acerca del método de evaluación del potencial.³ Señala que la evaluación del potencial debe realizarse exclusivamente a través de la línea ejecutivo-gerárquica de la organización elige esta modalidad a pesar de haber sido el autor, junto con Cason, del método de análisis del discurso, al que considera valido luego de un entrenamiento serio para la investigación o eventualmente para los procesos de selección que operan con ingresantes que no son conocidos por los gerentes.

Nosotros coincidimos con esta posición por considerar que el método desarrollado a través del gerente del gerente es más práctico y objetivo y no promueve ansiedades innecesarias que se producen cuando los empleados son evaluados desde afuera. Este método, lejos de alentar la incomodidad, generar ansiedades innecesarias y provocar sensaciones de invasión en lo personal, promueve la certeza de ser valorado y tenido en cuenta.

PROCEDIMIENTO PARA LA DETERMINACIÓN DEL POTENCIAL

Responsabilidades específicas de los roles que intervienen en el proceso de evaluación del potencial

1. la responsabilidad de la evaluación.

³ Comunicaciones personales, en seminarios realizados por el estudio Schlemenson y asociados, año 2000.

Como ya lo hemos establecido, la responsabilidad por la evaluación del potencial atañe al gerente de la población evaluada (g del g).

- evaluar es formular un *juicio gerencial*. El G del G es, según se deduce de estas definiciones, el responsable último por el juicio que se consigna, pero para llegar a una conclusión debe confrontar su juicio con el de sus subordinados inmediatos. La necesidad de llegar a una sola conclusión promueve un rico proceso de negociación de juicios que se constituye en un control intersubjetivo y en una prueba de la consistencia de los mismos. La evaluación es una apreciación individual acerca de la capacidad potencial.
- El G del G es quien posee la distancia necesaria para realizar dicha apreciación. Es, por otra parte, el principal responsable por la cobertura de cargos de su área amplia y, por lo tanto, es quien necesita prever los reemplazos eventuales de sus empleados directos y las coberturas de los puestos que dependen de él.
- El trabajo de evaluar simplifica y es claramente entendido por los gerentes cuando los juicios se refieren a los niveles o puestos ubicados en un estrato determinado.
- El *lenguaje sintético de los resultados* ofrece una ventaja para la sistematización de la información y el armado de una base de datos. El sistema que presentamos se expresa en una conclusión muy sintética, práctica y contundente, que en el lenguaje de los estratos y de las bandas para cada estrato se formula con un número y una letra. Antes definimos operacionalmente el potencial como el nivel más alto que alguien puede desarrollar hoy. En el lenguaje de los estratos, el número romano expresa el nivel del estrato y la letra expresa la banda dentro del estrato a); por alto; m: por medio; b: por bajo). De manera que el resultado en un listado con todos los candidatos que forman parte de un grupo objetivo, al lado de cuyos nombres se anota el estrato y la banda.

2. la responsabilidad por la aplicación del sistema en toda la organización.

El ejecutivo principal de una organización en la que se desarrolla un sistema de evaluación del potencial de esta naturaleza es quien tiene la *responsabilidad* por:

- decidir la implementación de la evaluación;
- liderar el balance correspondiente del *pool* de talentos;
- comenzar el trabajo de desarrollo de carrera con los colaboradores de sus colaboradores inmediatos, actividad que marca el comienzo del proceso de *mentoring* a través y a lo largo de la organización. El mencionado gerente es quien puede asegurar la implementación homogénea del sistema, que comienza con la formulación correspondiente de las políticas que avalan los objetivos.

No siempre el ejecutivo principal de una compañía o de un organismo público aprecia la importancia de las políticas específicas de recursos humanos relacionadas con los temas que venimos desarrollando y que son de su particular incumbencia. La política de recursos humanos de evaluación del potencial debiera ser una cuestión de conducción superior que se trasmite en cascada al resto de la organización. Una resolución de esta naturaleza merece ser comunicada por escrito a todos los miembros de la organización, y de esta forma se incorpora la cultura y a las prácticas internas derivadas de ella. En realidad, la única forma

de cohesionar la organización y motivar a la gente surge de una política coherente de recursos humanos asumida por el equipo de conducción superior con la asistencia y el liderazgo profesional del área de Recursos Humanos. Cuando esta premisa se cumple, se convierte en el principal factor de aliento para la implementación de un sistema.

3. responsabilidad del área de Recursos Humanos:

El área de Recursos Humanos:

- es responsable por la definición, implementación, monitoreo y auditoría del sistema de evaluación del potencial;
- se hace depositaria y custodia la información de las evaluaciones de toda la organización que forman parte de la base de datos de evaluación del potencial, manteniendo la confidencialidad de la información reunida;
- promueve la utilización de la información recogida y ofrece información a las áreas cada vez que se tienen que tomar decisiones de promoción, establecer reemplazos o designar sucesores; ofrece, en tal sentido, un panorama global amplio que comprende a la totalidad de la organización.
- Asegura el mantenimiento actualizado de la base del potencial.

Toda la información resultante de la evaluación del potencial tiene carácter confidencial y su circulación es restringida. La conoce el G del G que es el responsable directo de su área. A su vez, la concentra y custodia el área de Recursos Humanos que protege la confidencialidad de la base.

4. Responsabilidad específica de los consultores externos

En primera instancia, los consultores consustanciados con la práctica, la metodología y la teoría en la que esta propuesta se apoya, tienen la misión de transmitir la propuesta, desde el punto de vista conceptual y operativo, al conjunto de la organización cuando ésta es aprobada y se los autoriza para hacerlo. La totalidad del proceso que implica un proyecto de esta naturaleza involucra una asociación directa con el área de Recursos Humanos que tienen la responsabilidad de revisar la tarea.

Específicamente, el equipo de especialistas desarrolla las siguientes acciones:

- Coordina los ejercicios de evaluación del potencial.
- De asesoramiento y asistencia los gerentes en todo proceso. Esto incluye la exposición y explicación del marco conceptual y el soporte técnico para facilitar la toma de decisiones.
- Asegura la correcta aplicación del método adoptado.
- Confecciona el material impreso requerido para el proceso de evaluación del potencial. Incluye el listado del grupo objetivo que está siendo evaluado.
- Asiste y monitorea las reuniones de equilibración.
- Puede acompañar al G del G en el proceso de devolución y monitoreo, en especial cuando no posee la experiencia necesaria.
- Elabora un informe final que incluye el análisis de la información, los resultados y las recomendaciones.

Procedimiento para la evaluación del potencial

1. Etapa de presentación e inducción del proyecto.

La introducción e inducción de un proyecto de evaluación del potencial constituye una etapa muy importante. Para cumplir con ella, realizamos diversas reuniones de presentación del proyecto, dirigidas a distintos grupos de funcionarios.

En la primera etapa se reúne en particular a todos los G de G de la población objetivo que dependen en forma directa del ejecutivo principal. Este grupo conforma la conducción superior. Se realizan seminarios introductorios con el siguiente esquema:

- a) Presentación del proyecto y sus objetivos. Explicitación de los conceptos de capacidad potencial actual, capacidad potencial futura.
- b) Explicitación de las políticas de RR. HH. Que guían el proyecto.
- c) Ventajas, para la organización y para los individuos involucrados en la población objetivo, de participar en un proyecto de esta naturaleza.
- d) Detalle del procedimiento de evaluación del potencial.
- e) Compromiso de devolución y comienzo del proceso de desarrollo de carrera individual.

2. Identificación de la población a ser evaluada. Identificación del G del G y del G *directo*.

La evaluación del potencial se realiza por etapas y los estratos se van cubriendo e a uno. Para cada etapa se necesita identificar el grupo objetivo, el G del G y el G directo. Como se expresó anteriormente, es fundamental que ambos gerentes tengan un conocimiento adecuado de sus subordinados directos e indirectos para estar en condiciones de emitir un juicio acerca del potencial de cada uno de ellos. De manera que, si este conocimiento no está dado, es necesario planear un programa de entrevistas para que el G del G conozca cada uno de los miembros del grupo que le corresponde evaluar en su potencial, lo que constituye un valor agregado del programa, pues fortalece su liderazgo. Este conocimiento se traduce en una transparencia mayor del área en cuestión y en una disminución del anonimato, lo que motiva a la gente e incrementa la pertenencia.

3. El G del G evalúa a su grupo de empleados o colaboradores mediatos.

La actividad siguiente consiste en una reunión del consultor de un miembro del equipo de recursos humanos con el G del G primero, para asistirlo en la evaluación y formulación del potencial de todos sus subordinados o colaboradores indirectos. Para ello, deberá asegurarse que el evaluador conoce y comprende la correspondencia entre los estratos organizacionales y los roles de los individuos que están siendo evaluados. Es menester, además, repasar los conceptos fundamentales, para hacer foco en lo que efectivamente queremos evaluar; se hace necesario distinguir entre capacidad actual, capacidad potencial y capacidad potencial futuro. El evaluador expresa sus juicios en forma rápida y global utilizando el lenguaje de los estratos, y la división entre alto, medio y bajo de cada estrato.

4. El G directo evalúa a su grupo de colaboradores inmediatos

Una vez cumplía la etapa anterior, se entrevista al G directo del grupo que está siendo evaluado. El procedimiento es exactamente el mismo y la información recogida se consigna en una columna lateral y paralela a la que se completó con el G del G.

5. Comparación de juicios evaluativos y análisis de las discrepancias.

Suelen darse casos en que, como corolario de las dos etapas anteriores, los juicios del G del G y del G directo presentan discrepancias significativas. Se entiende por discrepancia a las diferencias de dos o más sub niveles entre los juicios de ambos. Por ejemplo, cuando uno hace “II alto” y el otro “II bajo”, o cuando no hay dos sub niveles de deferencia pero las evaluaciones se sitúan en estratos distintos. Por ejemplo, II alto y III bajo.

En los casos de discrepancias, el consultor se las señala al G del G y lo invita a la reconsideración del juicio evaluativo a la luz de esta información. Éste puede consultar al G directo o puede resolver la discrepancia por sí mismo. Dado que la decisión final y la responsabilidad la tiene el G del G, a él le corresponde en última instancia definir la discrepancia.

6. Graficación de las evaluaciones

Todas las evaluaciones reveladas hasta el momento se registran en una planilla de clasificación que refleja la capacidad potencial en estratos y en niveles dentro de los estratos.

Este material facilita la visualización de los resultados y permite eventualmente detectar posibles ajustes (procesos de equilibración). La graficación de los juicios forma parte de las tareas que desarrolla el equipo consultor.

7. Reunión de equilibración

En una reunión especialmente convocada, se encuentran en el G del G y todos sus gerentes subordinados o colaboradores directos, de quienes a su vez dependen los candidatos del grupo total que fue evaluado. En la reunión participa el consultor externo que actúa como asesor, con la colaboración de asistentes u observadores, quien facilita la graficación de las evaluaciones y consigna los cambios.

La información acerca de las evaluaciones realizadas en las etapas previas es presentada por el consultor en un gráfico de estratos y subniveles dentro de los estratos (alto-medio-bajo). En el gráfico figuran, con su nombre, la totalidad de los evaluados que dependen indirectamente del mismo G del G.

La ubicación de la población en un mismo gráfico invita a las comparaciones colaterales, lo que constituye un objetivo específico de la reunión. Esta instancia del proceso cumple con un doble propósito:

- Analizar y definir las posibles diferencias entre los juicios del G del G y de cada gerente participante en la reunión, es decir, volver a considerar las posibles discrepancias que subsistan.
- Someter a prueba de consistencia de los juicios evaluativos como resultado de las comparaciones colaterales y producir todos los cambios necesarios en las

posiciones relativas hasta que los participantes lleguen a un total convencimiento acerca de las evaluaciones. El gráfico ayuda a realizar las comparaciones porque permite visualizar, por ejemplo, todos los roles ubicados en una misma banda dentro del estrato y la relación que mantienen entre sí, hecho que promueve reubicaciones parciales hasta que se llega a la conformidad.

8. Procesamiento de los datos e informe final

En gabinete, el equipo consultor procesa la información y realiza un informe final dirigido al G del G. esta información se presenta bajo la forma de gráficos que tienen un carácter sintético.

9. Reuniones tutoriales y desarrollo individual: el proceso de mentoring

Las reuniones tutoriales constituyen un proceso por el cual el g del G, trabajando conjuntamente con sus gerentes inmediatos, ayuda a los colaboradores mediatos (subordinados de los subordinados) a comprender su capacidad potencial actual. A través de la entrevista tutorial, cada uno de los evaluados recibe un informe acerca del resultado obtenido y es ayudado por el G del G a conocer y entender su potencial para trabajar a partir de ello en su desarrollo individual laboral. En este proceso se tienen en cuenta las necesidades de la organización y las oportunidades disponibles. Esta información, correctamente suministrada, constituye un aspecto muy importante de la responsabilidad del G del G en relación con sus colaboradores indirectos. Deberá intercambiar con ello y hacerles conocer sus juicios evaluativos, y trabajar conjuntamente para determinar las actividades de desarrollo fuera de las que puede desplegar en su rol actual. El G del G tiene la responsabilidad de las que de saber si su interlocutor en un posible sucesor de un subordinado inmediato y si está preparado para otras asignaciones.

Por medio de la reunión tutorial:

- Se brinda la oportunidad a los colaboradores mediatos de explicitar sus intereses laborales, valores y aspiraciones, dentro del contexto de las necesidades de corto, mediano y largo plazo de la organización;
- Se les ayuda a comprender su potencial y a desplegarlo en toda su amplitud;
- Se discute con ellos acerca de lo que deberían hacer para desarrollarse en una perspectiva más amplia que la que ofrece su rol actual;
- Se exploran oportunidades de desarrollo, experiencias, transferencias o rotaciones, asignaciones a proyectos especiales, posibles ascensos, etcétera.

10. Tiempos para las reuniones tutoriales

Las reuniones tutoriales deberían realizarse al menos una vez por año. El G del F deberá organizar un cronograma formal que incluya a todos los subordinados de sus subordinados a fin de asegurarse de que los planes de desarrollo que implementan contemplan las aspiraciones del empleado, y de que cubren tanto las necesidades del individuo como las de la organización.

Los empleados tienen el derecho de conocer sus evaluaciones. Si no se las da a conocer, suelen utilizar mucho tiempo tratando de hacerse una idea acerca de o que la organización

ha planeado para ellos. La información que obtienen por su cuenta muchas veces es incorrecta y, por esa razón, perjudicial. Este mecanismo constituye un recurso fundamental para mejorar las comunicaciones y para que los individuos se sientan inmersos en un plan de desarrollo de carrera.

Para los interesados, el hecho de ser identificados con la capacidad necesaria para llevar a cabo tareas más complejas es muy importante y constituye de por sí un reconocimiento. Para los gerentes, una comunicación clara y concisa dirigida a los empleados refuerza el vínculo; el acto de emitir tal comunicación genera un mayor conocimiento e información acerca del *pool* de talentos de la organización y fortalece el desarrollo gerencial.

Los empleados deben tener la oportunidad de responder, de decir qué piensan acerca de las evaluaciones de su potencial. Además, por este medio comenzarán a asumir la responsabilidad por su propio plan de desarrollo. De esta forma, incrementarán su poder de elección y de decisión sobre su futuro laboral. Las recomendaciones del G del G se agregan, pero no reemplazan, a los propios planes del empleado con respecto a su crecimiento personal y carreras posibles.

DESCRIPCIÓN DE LAS ETAPAS DE UN PROYECTO EN GRAN ESCALA

Los datos que se exponen a continuación están referidos a un caso real, y sólo se han introducido cambios en las denominaciones y en los resultados para proteger la confidencialidad de la información recopilada.

Primera etapa

El grupo objetivo inicial está conformado por los subordinados de los gerentes especializados de áreas, que dependen en forma directa del director general, el ejecutivo principal de la organización.

En una organización grande, esta etapa puede comprender de 150 a 300 evaluaciones según la amplitud del control que ejerce cada uno de los gerentes de áreas, titulares de las principales áreas especializadas de la organización. Por tratarse de una gran corporación, la etapa comprende al director general, ejecutivo principal de estrato VI, que evalúa conjuntamente con sus gerentes de área de estrato V a los gerentes especializados de estrato IV.

En esta primera etapa el ejecutivo principal es el G del G del grupo objetivo de estrato IV.

Segunda etapa

Evaluación del potencial de los gerentes y especialistas correspondientes al estrato III. Esta etapa corresponde a unidades operativas, gerentes operativos de áreas y puestos equivalentes de áreas especialistas (asesores y marketing, sistemas y otros), descentralizadas, autónomas, y puede comprender la evaluación de alrededor de 1000 personas que ocupan roles en el estrato mencionado.

Tercera etapa

Evaluación de la capacidad potencial de gerentes de estrato II, que completa la etapa anterior. Se trata de posiciones de gerencia de primera línea de estrato II, entre analistas, profesionales, etcétera.

En síntesis, un proyecto de esta naturaleza, que comprende un número de más de 3.000 personas evaluadas en todos los niveles, puede ser realizado con comodidad a lo largo de un período de uno o dos años, dependiendo del tiempo de dedicación y de la intensidad que se destinen a la tarea.

ALGUNAS CONCLUSIONES DEL ANÁLISIS REALIZADO

En términos muy globales, el análisis puede revelar, por ejemplo:

- Abundancia de candidatos con posibilidades de ocupar cargos de estrato III, que actualmente están ubicados en posiciones de estrato II;
- Un número menor de candidatos con posibilidades de ocupar cargos del estrato IV, que actualmente ocupan posiciones de estrato III y algunos de estrato II;
- Por encima del estrato IV, se advierte una gran escasez de candidatos con potencial adecuado para ocupar posiciones de estrato N o más;
- La cantidad de empleados con un potencial superior a los cargos que ocupan excede las necesidades de cobertura actuales, lo cual revela que hay un número significativo de personas que están trabajando por debajo de su nivel de CPA y reciben una remuneración inferior a la que les correspondería si pudieran ejercerla;
- Hay un % de empleados que están trabajando, desde el punto de vista de su potencial, en un estrato inferior al de sus puestos actuales y el desajuste puede ser gratificado.

Datos globales

De las evaluaciones del total de empleados surgen los siguientes datos: el 76% trabaja a nivel del rol. Su CPA se encuentra en equilibrio con el nivel de puesto expresado en términos de estratos. El 18% presenta un CPA superior a su rol. De aquí surge el posible *pool de talentos*, en el cual la empresa podrá poner la mira para desarrollar sus cuadros de reemplazo e intensificar el desarrollo de carrera. Suele haber, dentro de este grupo, marcadas diferencias; un residuo número de individuos, ubicados en las *curvas de pregestión de la capacidad potencial*, podría llegar a ocupar posiciones estratégicas, equivalentes a las gerencias de áreas especialistas.

Por último, el 6% ha sido evaluado por debajo de su rol. Estos individuos demuestran estar por debajo de las expectativas actuales de sus jefes. La empresa deberá tomar alguna determinación en relación con ellos.

CONCLUSIONES GENERALES SOBRE EL PROYECTO, LA METODOLOGÍA Y SU IMPLEMENTACIÓN

Acerca del proceso y del método de implementación

Todas las personas involucradas en el proceso y, en particular, las que formaron parte del grupo objetivo, reciben amplia información acerca del proyecto: políticas, conceptos, principios, metodología, criterios de confidencialidad, beneficios, etcétera. Los seminarios que se realizan tienen por finalidad brindar esa información.

La transparencia intencional que supone este criterio favorece un vínculo de confianza con el proyecto y con el equipo, fundamental para lograr un clima de colaboración y compromiso por parte de los miembros de la organización.

Toda vez que se cumple este criterio, los participantes demuestran un buen grado de valoración y deseo de participación, especialmente cuando el proyecto es avalado en forma activa y comprometida por el ejecutivo principal de la organización. Lejos de despertar temores, los comentarios realizados por los participantes revelan preponderadamente interés y seriedad, y el deseo de involucrarse en un proceso que e ocupa seriamente de considerar a la gente.

Los gerentes de los niveles superiores toman estas presentaciones como verdaderas oportunidades de entretenimiento en temas que tienen que ver con la evaluación individual, la conducción de equipos de trabajo y el gerenciamiento. Perciben, con buen criterio, que el ejercicio propuesto contribuye al fortalecimiento de la autoridad del gerente y refuerza, en este caso, el liderazgo del G del G que, en la mayoría de las organizaciones, no se ejerce a pleno.

El interés inicial en el proyecto se mantiene en todas las etapas del mismo.

Algunas condiciones que hacen a la efectividad y los buenos resultados

Algunas de las conclusiones que se expresan a continuación ya fueron mencionadas en el desarrollo de este capítulo, no obstante consideramos necesario volver sobre ellas.

- Para que el proceso se pueda desarrollar en forma rápida y sencilla, es menester que el G del G conozca a sus subordinados indirectos.

La evaluación propiamente dicha se vuelve fluida cuando el G del G y el G directo conocen a los evaluados. Este conocimiento se profundiza, en el caso de los gerentes, a través del acto de delegación de trabajo, cuando: a) les fijan objetivos en términos de “qué”, “cuanto” y “para cuándo”; b) supervisan el trabajo; c) realizan la práctica de *coaching*. El la medida en que el gerente es evaluado por los resultados de sus colaboradores inmediatos, el compromiso entre ellos es fuerte y este hecho profundiza el conocimiento recíproco.

El conocimiento del G del G se profundiza a través de las reuniones de mentoring. Este conocimiento directo del grupo ampliado que compone el sector y que está conformado por los colaboradores de los colaboradores directos del G del G, contribuye a una atmósfera social de transparencia, que permite saber *quién es quién* y que se profundiza por el conocimiento recíproco. Esta instancia se denomina *factor de mutuo reconocimiento*, y favorece la personalización del sistema que reduce el anonimato y la alienación.

Un G de G puede tener entre setenta y noventa subordinados indirectos, o a veces más. De manera que el conocimiento previo de sus colaboradores *mediatos* constituye una tarea que le insumirá tiempo. Es un tiempo sumamente productivo que reforzará su liderazgo en el sector al ofrecerles a sus colaboradores indirectos la oportunidad de fortalecer su identificación con dicha figura y con los objetivos del área. Podrán apreciar su forma de pensar, su personalidad, sus juicios acerca de la carrera individual de cada uno, y todo ello tiene un efecto de motivación e interés.

La repetición de este procedimiento dará lugar a las reuniones tutoriales p de mentoring, a las que antes hemos hecho referencia y que constituyen una oportunidad de aprendizaje para el desarrollo de carrera.

Al finalizar el procedimiento, el G del G tendrá un conocimiento mucho más personalizado de la gente de su área y se habrá profundizado, así, una cualidad a la efectividad y confiabilidad de las conclusiones.

Hacer foco significa centrarse en aquello específico que se evalúa. Habrá que cerciorarse de que los interlocutores entienden en forma precisa qué es *capacidad potencial*, lo cual insume una buena porción de las entrevistas iniciales. Entender esta noción es fundamental, para lo que hay que diferenciar los tres conceptos vinculados de capacidad. Cuando los gerentes hacen foco en lo que efectivamente se evalúa, recuperan una vivencia enriquecida del conocimiento de sus empleados directos y pueden manifestar con mucha precisión lo que se les pide. Formulan sus juicios globales y su apreciación sobre el nivel más alto en el que pueden trabajar sus colaboradores indirectos o subordinados de sus subordinados y lo hacen con una notable rapidez y precisión.

- La evaluación es referenciada al nivel y al puesto.

Como ya señalamos en distintas oportunidades, el ejercicio requiere una apreciación muy clara del nivel del puesto que se está evaluando. Con este fin, se suministra un documento de trabajo que consiste en un gráfico de los niveles organizacionales de los lapsos temporales.

En el gráfico estarán colocados los puestos principales en los niveles y señaladas las situaciones de comprensión, cuando existen. Los jefes asimilan el lenguaje de los niveles y lo usan con precisión. Una vez que el entrevistado sintoniza el puesto con el nivel, pasa a hablar inmediatamente en términos de estratos y de las diferencias de cada uno de ellos.

La realización de esta etapa, al ejercicio propiamente dicho, ofrece como subproducto un diagnóstico de organización que es sumamente útil. De él se derivan una cantidad de propuestas para el diseño que muchas veces no forman parte de los objetivos iniciales de un proyecto de evaluación del potencial, pero que brindan un valor agregado que surge de la correcta aplicación del método.

- La reunión de equilibración sintetiza la esencia del proceso.

En la reunión de equilibración, los participantes recorren nivel por nivel las posiciones de los evaluados que están ubicados en dependencias diferentes, pero, dado que forman parte del área de un mismo G del G, es muy probable que los participantes conozcan a todos los evaluados. El G del G es el que se encuentra en mejores condiciones para realizar comparaciones, pues se supone que conoce y ha evaluado a todos los empleados. Cada

gerente es libre de manifestar sus juicios no sólo sobre sus colaboradores inmediatos, sino también sobre los de sus colegas, pudiendo realizar una defensa de sus opiniones. De hecho, como resultado de las comparaciones del potencial de unos con otros, comienzan a producirse traslados en las posiciones relativas que tienen a hacer más coherente la graficación total.

En la reunión pueden reconsiderarse nuevamente la totalidad de las evaluaciones previas, a la luz de las comparaciones colaterales. De esta forma se someten a prueba de conciencia los juicios evaluativos. Esta comparación entre pares permite lograr la sinfonía fina del proceso. Todos pueden opinar, pero el G del G es el que en última instancia decide, luego de dejarse asesorar y escuchar argumentos en uno u otro sentido.

Acorde con el procedimiento, el consultor invita a los participantes a recorrer visualmente el gráfico que aporta a la reunión comenzando por el nivel más alto. En esta oportunidad, los participantes explicitan sus criterios y analizan los casos más relevantes en una actividad que resulta ser sumamente constructiva. Se habla directamente de la gente, en participar de los casos menos claros, se hace alusión directa a las situaciones de trabajo en las que se pudieron observar rasgos de conducta relevantes, manifestaciones del talento, dificultades o inhibiciones. Cuando los jefes se enfrentan con alguna discrepancia, ponen de manifiesto que no están utilizando de forma correcta algunos de los conceptos en juego. A veces confunden el potencial con aspectos de la personalidad o con dificultades de integración al equipo de trabajo, o con aspectos que atañen la ética de los individuos, y utilizan estos antecedentes para bajarlos de nivel en las evaluaciones, cosa que no corresponde en una evaluación del potencial, aun cuando correspondería en una evaluación de la efectividad. A su vez, discriminan con mucha lucidez cuándo un individuo sabe venderse bien, pero, a la hora de los hechos, no concreta o no puede solucionar problemas relevantes de su tarea y necesita recurrir en exceso a su gerente. La convicción de estar actuando con ecuanimidad y de ocuparse seriamente de la gente al evaluar su potencial trae como corolario un sentimiento constructivo y una satisfacción por la tarea realizada.

- Beneficios colaterales para la evaluación que surgen del ejercicio del *pool* de talentos.

Los resultados de cada una de las etapas del proceso son expuestos en informes que contienen datos específicos organizados en listados, gráficos, curvas de desarrollo y diagramas de discrepancia entre el nivel del puesto y el nivel de las personas.

La consistencia de los resultados que aportan las evaluaciones del potencial es susceptible de ser corroborada en la práctica por los G del G involucrados y or las propias figuras ejecutivas principales a través del conocimiento personal que poseen de los evaluados. Otros miembros de la organización también están en condiciones de corroborar la adecuación de los juicios emitidos a través de la trayectoria de las personas involucradas y de los aciertos o desaciertos que van teniendo.

La gente preparada para evaluar en el marco de un procedimiento sistemático como el que se detalla demuestra una gran sensatez. Al realizar la tarea con seriedad y responsabilidad, y al tener la posibilidad de corroborar sus juicios, chequeándolos con las opiniones de los demás se van acercando a apreciaciones sensatas y objetivas.

La realización del ejercicio del *pool* de talentos permite definir con precisión e perfil de los distintos estratos en términos del horizonte temporal, de la descripción d tareas y de su nivel de complejidad. A la vez, permite delimitar con precisión las competencias requeridas para ocupar dichos roles. En los casos en los que intervenimos, siempre hemos podido

aprecia con mucha claridad ciertos niveles organizacionales clave que debían ser reforzados o que aparecían como innecesarios, de esta forma el proyecto contribuyó a definir con claridad un diagnóstico de la organización.

BLOQUE IV

“EL PROYECTO DE EVALUACION”

H.S. Bhola. “Paradigmas y modelos de evaluación”. En: La evaluación de proyectos y campañas de alfabetización para el desarrollo. Instituto de la UNESCO para la educación. Antología UPN “Evaluación y seguimiento en la escuela LE’94” pp. 156-163. Ed. Santiago de Chile (1992 pp. 186-198)

LECTURA:

CAPITULO 2 DE PARADIGMAS Y MODELOS DE EVALUACION*

PRESENTACION

Los diversos modelos de evaluación son divididos en esta lectura en dos grupos, ubicados de acuerdo con los paradigmas que los gobiernan, esto es considerando sus posiciones filosóficas e ideológicas que los crearon. Estos dos paradigmas se han denominado: paradigma racionalista y paradigma naturalista.

Además encontraron que los diversos modelos de evaluación se orientan hacia cada uno de los paradigmas que se presentan.

Alguno de los modelos revelaron los comienzos de ideas que han llegado a ser parte de la sabiduría convencional. Otros mostraron claramente cómo se apuntalan en los métodos y las técnicas que ahora son de uso frecuente. Aún otros pueden sembrar las semillas de la duda acerca de las cosas que todos daban como ciertas.

Entre los académicos y entre quienes ejercen la práctica de la evaluación existen diferencias de la manera de ver el mundo, debido a que en sus experiencias profesionales se han puesto en práctica modelos distintos de evaluación.

Al tratar de usar estos modelos no tenemos que buscar utilizar uno de ellos en forma pura y completa, debido a que pueden ensayarse mezclas de esto, según sea el caso y el ámbito donde se evaluará.

PARADIGMAS Y MODELOS DE EVALUACION

¿Qué es un paradigma?

Comencemos con la palabra “paradigma”. Los significados que da el diccionario son pretenciosos. El paradigma se define como una lista ordenada, un cuadro de clases, un

* h. s. Bhola. “Paradigmas y modelos de evaluación”. en: La evaluación de proyectos, programas y campañas de alfabetización para el desarrollo. Instituto de la UNESCO para la educación. Editado en Santiago de Chile, 1992. Pp. 27, 52.

patrón o una fórmula que define el modo general en que deben colocarse los casos específicos de un cierto orden. En términos formales, se ha definido un paradigma como un sistema axiomático con un determinado conjunto de supuestos acerca, de los fenómenos sobre los cuales se pretende indagar.

Kuhn¹, en su estudio sobre las revoluciones científicas, definió de manera pintoresca un paradigma como la ideología creativa de los científicos, desde la cual ellos trabajan y la que les proporciona una posición lógica y metodológica particular que les permite producir conocimiento científico o socio-científico.

Así, los paradigmas de evaluación son las ideologías creativas de los evaluadores. Estos, paradigmas determinan el pensamiento y el comportamiento metodológico de los evaluadores: lo que ellos piensan sobre la naturaleza de la realidad; y cómo piensan que puede llegarse a establecer “afirmaciones fundamentadas”, esto es, declaraciones dignas de crédito acerca de la realidad social que nos rodea. Hay dos paradigmas básicos de evaluación (e investigación) que analizaremos: el paradigma racionalista y el naturalista.

¿Qué es un modelo?

Definiremos ahora modelo. Formalmente un modelo es un conjunto de información, datos o principios agrupados de manera verbal o gráfica (a veces en forma matemática) para representar o describir alguna cosa, idea, condición o fenómeno. En un lenguaje menos formal, un modelo es la esencia del aprendizaje y el pensamiento de un especialista presentado clara y brevemente. Los modelos son la prole de los paradigmas. Es importante recordar esta relación de padre a hijo entre los paradigmas y los modelos. Así, los modelos. Así, los modelos de evaluación pueden ser presentaciones verbales, en cuadros o gráficas, de los principios que han aprendido los evaluadores. Son la esencia de sus experiencias por separado desarrolladas en el marco de determinados paradigmas. En otras palabras, son un conjunto de valores, un conjunto de preferencias y un conjunto de procedimientos presentados como una unidad.

Finalmente, los evaluadores pueden hablar a veces de enfoques y de estrategias de evaluación. Estos no son más que métodos para comenzar o desarrollar un estudio, evaluativo. Presumiblemente, cuando un enfoque o una estrategia utilizada y sometida a prueba sucesivamente llega a estandarizarse y formalizarse, adquiere el status de modelo de evaluación.²

SECCION A: Dos Paradigmas Básicos de Evaluación

Hay dos paradigmas básicos de evaluación en la alfabetización, la capacitación y el desarrollo.

1. Evaluación Racionalista (ER) y
2. Evaluación Naturalista (EN)

Esta clasificación es completamente satisfactoria. El rótulo de evaluación racionalista, ER, se ha utilizado como sustituto de enfoques positivistas lógicos de la evaluación, aún cuando somos plenamente conscientes de que no comprende el conjunto total de supuestos del positivismo lógico. Y tampoco nos ayuda este rótulo a comprender que ahora se utilizan diferentes versiones del positivismo ni que esté surgiendo una versión reacondicionada del

positivismo lógico. Por otro lado, el rótulo de evaluación naturalista, EN, tampoco es muy satisfactorio. Sólo porque la EN se ofrece como alternativa a la evaluación racionalista, no debe concluirse que la naturalista es irracional. EN se emplea aquí como un término incluyente para un enfoque que es a la vez constructivo y colaborativo. Estos dos términos se explican más adelante. Aquí basta decir que ambos paradigmas –ER y EN- son “científicos” aunque se diferencien en los supuestos que asumen sobre “cómo funciona el mundo” y “qué y cómo podemos aprender acerca del mundo”.

El paradigma racionalista

Al paradigma racionalista también se le conoce como paradigma del positivismo lógico. Asume que la realidad existe “afuera” y cualquiera puede verla o experimentarla a través de los sentidos. En otras palabras, enfatiza lo explícito -lo que puede afinarse directamente y con certidumbre. Seguir este paradigma es tener una actitud “positiva” acerca de las afirmaciones que se hacen con respecto a la realidad, y depender de ser “lógicos” al derivar subsiguientes afirmaciones verdaderas.

Sus tres características esenciales son: el redaccionismo -que las partes pueden separarse del todo para estudiarlas sin que haya cambios: repetibilidad- que lo que alguien ha descubierto puede repetirlo otro; y refutación -que lo que se afirma debe poder confirmarse o refutarse. La gran esperanza de quienes siguen este paradigma es, desde luego, generar afirmaciones a manera de leyes, universalmente generalizables.

A veces se le llama paradigma clásico porque ha sido usado por mucho tiempo, sigue reglas estrictas y se considera -aunque esto no sea necesariamente correcto- como estándar y autorizado.

El paradigma racionalista sigue los métodos de las ciencias exactas algunas veces llamadas restringidas, como la física, la química y la ingeniería. Su ideal metodológico es usar muestras al azar y el experimento controlado. Bajo ciertas condiciones pueden ser aceptados los diseños cuasi-experimentales³. El paradigma racionalista demanda una clara definición de los objetivos de evaluación y de las variables, un plan de muestreo, una instrumentación estructurada que genera datos cuantitativos, técnicas estadísticas para el análisis de los datos y generalización de los resultados.

El paradigma naturalista

El paradigma naturalista asume que la realidad no existe afuera para que cada quien la vea y la experimente de la misma manera, sino que el mundo se encuentra (como realidad objetiva) y se elabora (esto es, cada individuo lo construye socialmente). En efecto, la parte más importante de nuestra realidad es construida socialmente. El evaluador o el investigador busca encontrar los significados que las personas llevan en sí. El paradigma naturalista sugiere que el comportamiento humano sea estudiado tal como ocurre naturalmente, en ambientes naturales y dentro de su contexto total. En otras palabras, el paradigma naturalista es de naturaleza holista por su orientación, y busca estudiar la realidad como un todo, sin dividirla artificialmente en partes y segmentos para ajustarla a la conveniencia del evaluador.

A veces se hace referencia a este paradigma como cualitativo y fenomenológico. Esto quiere decir que a diferencia del evaluador racionalista, el naturalista busca primero descubrir los fenómenos luego ir en busca de métodos y modelos. El evaluador naturalista

va en busca de comprender la situación específica que puede iluminar después otras situaciones similares. El evaluador naturalista no pretende obtener leyes generalizadas, sino ideas perspicaces que puedan transferirse de un contexto a otro.

En la indagación naturalista, los métodos empleados son los del antropólogo y el etnógrafo. El evaluador o investigador mismo constituye parte del fenómeno que estudia -no puede considerarse aislado objetivamente “fuera de” la realidad que estudia. El diseño de EN va surgiendo; aparece según el evaluador emprende diferentes pasos y sigue distintos procedimientos para recolectar datos significativos. Las muestras son propositivas más que aleatorias. Los instrumentos son siempre inestructurados generan datos cualitativos. Se busca la aplicabilidad y adecuación de los resultados más que su generalización.

El paradigma racionalista versus el paradigma naturalista

No es posible incluir en el ámbito de esta monografía un análisis más completo de estos dos paradigmas. En el cuadro hemos resumido las diferencias entre estos dos paradigmas de evaluación a fin de que el lector pueda examinarlas cuidadosamente.

Ahora se sugiere que muchos de los puntos fuertes que se atribuyen a paradigma racionalista son meras suposiciones: algunas de éstas no tienen defensa conceptualmente otras no pueden sustentarse en la práctica. Por ejemplo, con frecuencia es difícil seleccionar criterios apropiados para juzgar el mérito o el valor de los programas; y los evaluadores que usan el paradigma racionalista ocasionalmente terminan eligiendo criterios restringidos que se acomoden a sus planes experimentales. Las situaciones experimentales son en esencia incontrolables, a pesar de que se proclame lo contrario. Es casi imposible la selección aleatoria de las muestras. Los grupos que se estudian frecuentemente se eligen a sí mismos y son sistemáticamente diferentes unos de otros.

Es imposible estandarizar los tratamientos experimentales en diferente lugares. Nunca se está seguro sobre qué es lo que se está comparando. Hay limitaciones en cuanto a la información que puede producirse al emplear este paradigma. Su utilización no produce información sobre proceso ni sobre los efectos netos de intervenciones distintas. Realmente se genera mucha información que no es útil y está “muerta de antemano”.

Existen, desde luego, muchas dificultades administrativas en la utilización del paradigma

racionalista de evaluación. Las necesidades del programa y las de la evaluación racionalista

frecuentemente van hacia direcciones diferentes. La elección entre grupos que recibirán o

no recibirán tratamiento -y recursos para el desarrollo- no es un asunto fácil de decidir.

DIFERENCIAS ENTRE LOS DOS PARADIGMAS BASICOS DE EVALUACIÓN

El Paradigma Racionalista

El Paradigma Naturalista

Raíces filosóficas

Fenomenológico

Positivista

Holista

Reduccionista

Libre de valores

Cargado de valores

Orientación teórica

Usa teorías bien fundamentadas

Somete a prueba la teoría existente

Relaciones plausibles

Conformación mutua y simultánea

Relaciones causales

Diseño

El diseño surge (o se desenvuelve) asegurando resonancia sin separar a quien conoce de lo que se conoce

Experimental o causi-experimental para asegurar objetividad y validez

La situación de evaluación/Investigación

la Ecológica, en el contexto natural

En el laboratorio o controlada

Muestreo

Propósito, de élite, especializado

Aleatorio

El tamaño se determina en la marcha, se completa cuando la información disponible es completa

Tamaño predeterminado

Orientación metodológica

Orientación hacia los objetos

Cuantitativa

Libre de objetos

Cualitativa

Cargada de descripciones

Instrumentos

Estructurados, frecuentemente

No estructurados, con frecuencia que no perturban la situación natural

El evaluador o investigador mismo se convierte en una herramienta de recolección

intervencionistas

de datos.

Se busca que sean estandarizados e independientes de sesgos del evaluador

Todo conocimiento es aceptable

Preferencia por datos exactos

Análisis de los datos

Temáticos

Típicamente estadístico

Análisis de contenido de entrevistas, documentos y observaciones

Informe

Descriptivo, interpretativo

Estadístico-analítico

Típicamente un estudio de caso

Naturaleza de las declaraciones de Verdad

Intuiciones acerca de covariaciones naturales de los acontecimientos

Leyes generalizadas

Ideas perspicaces, analogías

Resultados convergentes conducentes a hacer predicciones

Resultados divergentes

Realidad tangible única

Realidades múltiples o una construcción negociada de la realidad en su contexto

Puntos fuertes

Es responsivo, el énfasis es holista y humaniza la actividad evaluadora

Proporciona buenas estimaciones de las diferencias, variaciones y correlaciones cuando las variables, realmente, pueden ser definidas apropiadamente y pueden establecerse controles razonables

Puntos débiles

El evaluador puede perderse en las complejidades de la vida real, puede carecer de habilidades interpersonales y perceptibilidad individual, por lo cual puede terminar haciendo afirmaciones poco significativas, basadas en meras impresiones.

El buscar acomodar las preguntas de evaluación a los métodos y modos de análisis aceptables puede llevar a escoger preguntas triviales y artificiales y obtener resultados triviales e inútiles

Las cuestiones sobre confiabilidad y validez

Con demasiada frecuencia las discusiones sobre la evaluación racionalista (ER) y la naturalista (EN) dan lugar a discusiones sobre su validez y confiabilidad ¿Son objetivos los datos recogidos? ¿Son válidos? ¿Son confiables? Ahora se acepta ampliamente que hacer estas preguntas con estas palabras es juzgar la EN contra los estándares de la ER. Desde un punto de vista preciso esto no es justo, y desde el nivel conceptual es absurdo.

Lincoln y Guba⁴, en su libro más reciente, han analizado los diversos aspectos del rigor tal como se aplican a los dos paradigmas básicos de la evaluación. Ellos plantean que tanto el

racionalista como el naturalista pueden ser científicos y riguroso, en sus propios términos. Ellos sugieren que los términos apropiados para juzgar que tan buena y confiable es la EN son diferentes de los que se usan para la ER, como se muestra en el siguiente cuadro:

**TERMINOS RACIONALISTAS Y NATURALISTAS APROPIADOS
PARA DISTINTOS ASPECTOS DEL RIGOR**

Aspecto	Término Racionalista	Término Naturalista
Valor de verdad	Validez interna	Credibilidad
Aplicabilidad	Validez externa/generalización	Adecuación
Consistencia	Confiabilidad	Verificabilidad
Neutralidad	Objetividad	Confirmabilidad

¿Qué paradigma deben elegir los elevadores de programas de desarrollo, educación y capacitación?

El paradigma racionalista ha logrado grandes victorias en las ciencias exactas. Ha producido investigaciones que han borrado enfermedades de la faz de la tierra y ha puesto al hombre en la luna.

Tanto éxito obtuvo este paradigma que los científicos sociales (sociólogos, psicólogos, economistas, educadores y aún antropólogos) quisieron imitar el paradigma “científico” del físico y el químico. Usaron el paradigma racionalista como venganza ¡Les hizo sentir realmente científicos! Por muchos años aprendieron el método racionalista y lo enseñaron en la mayoría de los departamentos de ciencias sociales de las universidades.

Durante los últimos veinte años se ha llegado a caer en la cuenta de que el paradigma racionalista ha llevado a los científicos sociales a sentirse bien pero no necesariamente a obtener buenos resultados. Hemos descubierto que demasiadas veces la vida social no encaja en el modelo experimental. Al tratar de controlar las variables, segmentamos el comportamiento humano de una manera que no es natural y realmente cambiamos la naturaleza misma del comportamiento humano que se estudia. La agregación de puntales y el tratamiento estadístico de los datos puede parecer elegante e impresionar favorablemente, pero los resultados han sido triviales y engañosos. En el comportamiento humano el contexto es importante. No tenemos que estudiar el comportamiento solamente, sino el comportamiento en-contexto.

La mayoría de las veces el paradigma naturalista resulta más apropiado para el estudio del comportamiento humano. Después de haber sido rechazado como subjetivo y cualitativo, está siendo cada vez más aceptado. Según su metodología se vaya clarificando y se avance en las técnicas de análisis de los datos, el paradigma naturalista de evaluación encontrará un lugar adecuado en la metodología de la evaluación.

¿Deberá siempre utilizar el evaluador el paradigma naturalista para la evaluación del impacto amplio de los programas de desarrollo, alfabetización de adultos y capacitación para el desarrollo, alfabetización de adultos y capacitación para el desarrollo? La respuesta es: si no siempre, casi siempre. El consejo que ha dado Lee J. Cronbach parece ser útil. Cronbach⁵ señala que en el mundo de la educación y el cambio social, uno se encuentra con dos contextos diferentes: el contexto del control, en el que el evaluador no puede

controlar la situación social para ajustarla a las necesidades de la evaluación (por ejemplo: el estudio de los movimientos de los ojos al mirar un afiche instructivo de tamaño grande); y el contexto de la acomodación en el cual el evaluador no puede controlar la situación social para ajustarse a las necesidades de la evaluación sino acomodarse él mismo a las realidades existentes (por ejemplo: el estudio de la frustración la agresión entre los niños en un parque de juegos). En los contextos que se pueden controlar, que tal vez no sean muchos, es correcto usar el paradigma racionalista. En los contextos de acomodación, sin embargo, tendría más sentido el paradigma naturalista porque, por definición, este paradigma no busca perturbar las realidades como naturalmente existen.

SECCIÓN B: Modelos de Evaluación

Sobre el trasfondo de estos dos paradigmas generales. los especialistas han propuesto muchos modelos diferentes de evaluación. Pero ¿por qué hay tantos modelos de evaluación? ¿No hay una sola forma correcta de hacer la evaluación?

Antes, en este capítulo definimos un modelo como lo esencial del aprendizaje y el conocimiento de un especialista, establecido de manera clara y parsimoniosa para comunicarlo entre los profesionales y los que ejercen la práctica. Hay muchos modelos de evaluación diferentes debido a que los distintos especialistas han pasado por experiencias diferentes al aprender y hacer la evaluación y al reflexionar sobre sus experiencias han utilizado diferentes puntos de vista y valores.

Los modelos de evaluación también difieren porque han surgido en diferentes situaciones en que se desenvuelven los programas: dentro de la educación formal o dentro de situaciones de educación extra-escolar y no formal; y en situaciones de salud mental en un país industrializado o dentro de la educación para la vida familiar en el contexto de un país en desarrollo.

Finalmente, y lo más importante es que los modelos de evaluación son diferentes porque los especialistas han introducido consideraciones de valor adicional en su elección inicial de los paradigmas. Algunos modelos de evaluación enfatizan visión más sinóptica de la evaluación, sugiriendo que no sólo evaluemos el comportamiento de los llamados beneficiarios sino también el nuestro. Algunos modelos de evaluación sugieren que se emplee la imaginación en las evaluaciones de modo que no dependamos sólo de fríos cálculos. También los hay que sugieren que las consecuencias no previstas de las acciones de los programas pueden ser tan importantes como las que se esperaban y se anticiparon. Por lo tanto el modelo evaluación debe poder acomodar tanto las consecuencias anticipadas como las que no se esperaban. Algunos proponen que la evaluación sea realizada como una defensa y confrontación. Otros sugieren la evaluación participativa, en la cual tanto fines como los medios son determinados participativamente por todos los implicados organizadores, profesionales y beneficiarios.

Se puede observar claramente que los valores dirigen el desarrollo de los modelos de evaluación durante los últimos veinte años: (1) hay una utilización exclusiva o complementaria de las estrategias naturalistas; y (2) existe un movimiento hacia la inclusión de los beneficiarios de los programas en el diseño y la ejecución de las evaluaciones. Las palabras claves son holista y participativa.

A continuación se analizarán algunos de los modelos de evaluación a los que se hace frecuente referencia en la literatura. Las discusiones serán breves. En este libro incluimos esta discusión por dos razones: una educativa y otra política. El agente de desarrollo o de la

alfabetización debe tener alguna idea de que modelos de evaluación existen y cuáles son sus características. Esto es para su formación. Pero también hay una razón de tipo político. Un agente alfabetizador también debe poder justificar su elección de uno o varios modelos y debe poder sustentarla frente a un especialista externo. ¡No debemos permitir que los técnicos especialistas nos intimiden con el uso de nombres y frases poco familiares! Se discutirán los siguientes modelos brevemente:

1. El modelo orientado por objetivos de Tyler
2. El modelo de experimentación social
3. El modelo CIPP y la cuadrícula EIPOL
4. Las dos caras de la evaluación
5. La evaluación resporsiva
6. El modelo de evaluación discrepante
7. La evaluación transaccional
8. La evaluación libre de metas
9. Los enfoques investigativos
10. La evaluación iluminativa
11. La evaluación por expertos
12. El modelo de la evaluación como defensa
13. El modelo de evaluación participativa, y
14. El modelo de evaluación como estrategia para la situación específica (3-S).

1. Evaluación orientada por los objetivos

Este modelo se asocia al nombre de Ralph Tyler y es tal vez el más antiguo de los modelos de evaluación disponibles.

La evaluación que se realiza bajo este modelo busca hacer comparaciones entre los “resultados esperados” y los “resultados reales”. En otras palabras, los niños o adultos participantes en un programa o proyecto se examinan para ver si se han logrado los objetivos referidos a la adquisición de formas particulares de pensar, sentir y actuar. En términos prácticos, la evaluación se equipara con las pruebas.

Hay puntos importantes en este enfoque, que se centra en los resultados, un concepto que se en tiende con facilidad. No hay necesidad de definir grupos experimentales y de control, con lo que generalmente se desorganiza la rutina diaria en las escuelas y en las comunidades y cuya puesta en marcha puede ser costosa. Las mediciones reflejan objetivos claramente establecidos, por lo cual la confiabilidad no es una gran preocupación. Mientras que las pruebas inicialmente son referidas a criterios, pueden adquirir funciones con referencia a normas si se hacen comparaciones entre diferentes situaciones consistentemente.

Sin embargo, este modelo tiene serias desventajas. La información generada por las pruebas es demasiado estrecha para constituir una base sólida y amplia para juzgar el mérito o valor del programa total. La información generada por el modelo es terminal. Da pie a que su utilización directa para mejorar el programa sea muy poca.

2. El modelo de experimentación social

Este es un modelo que busca experimentar con grupos sociales ya existentes. La sociedad se convierte en laboratorio.

En la modalidad experimental clásica, el evaluador que usa este modelo elige dos grupos. Uno que recibe el tratamiento experimental y otro que no lo recibe. Los conceptos metodológicos esenciales son: distribución aleatoria, control, tratamiento y comparación.

Los proponentes del modelo han sugerido también los que se conocen como diseños cuasi-experimentales, los cuales se supone que se ajustan mejor a las realidades del mundo real. Sin embargo, recientemente sus mismos proponentes han planteado serias dudas con respecto a los diseños cuasi-experimentales.

3. El modelo Contexto- Insumos-Procesos- Producto (CIPP) y la red EIPOL⁶

El modelo CIPP se asocia con frecuencia con el nombre de Daniel L. Stufflebeam, quien lo ha usado en distintos estudios evaluativos.

De acuerdo con este modelo, el único propósito de la evaluación es producir información útil para las decisiones. Empleando la metáfora del sistema y su cuatro parámetros (contexto, insumos, procesos, productos), el modelo habla de cuatro tipos de evaluación para proporcionar información referida a cuatro tipos de decisiones:

1. Evaluación de contexto - para proporcionar información sobre la situación y poder tomar decisiones de planificación.
2. Evaluación de los insumos - para tomar decisiones de programación como diseño de proyectos alternativos y decisiones sobre personal.
3. Evaluación de los procesos - para tomar decisiones relacionadas con la metodología y la ejecución, y
4. Evaluación de los productos - para evaluar el impacto y tomar decisiones sobre reciclaje.

El modelo CIPP, cuando se propuso por primera vez, combinaba el vocabulario de sistemas con la investigación formal, con su énfasis en la clarificación de las necesidades de decisión en la evaluación, la observación estructurada y la tradición de aplicar pruebas de rendimiento en las escuelas. El modelo adoptó los criterios de validez interna y externa, confiabilidad, objetividad, pertinencia, importancia, alcance, credibilidad, oportunidad, permeabilidad y eficiencia de la información evaluativa producida. Se le ha criticado por mostrar poca preocupación por los valores. Algunas versiones recientes del modelo han tratado de superar algunas de estas críticas.

En la cuadrícula EIPOL, Ravinda H. Dave⁷ dividió el “resultado” del lenguaje de sistemas (llamado el “producto” en el modelo CIPP) en dos partes: (i) los resultados del aprendizaje y otros resultados “intermediarios” del programa y (ii) los efectos a largo plazo del programa en los dominios educativo y socio-económico. Así los cuatro parámetros del sistema (contexto, insumos, procesos, productos) se convierten en cinco dimensiones o fases de la evaluación, que son: la de la situación ambiental, la de los insumos, la de los procesos, la de los resultados inmediatos y la de los efectos a largo plazo. Combinando estas dimensiones con las cuatro fases principales del ciclo de un proyecto (pre-planificación, ejecución, asimilación) se genera la llamada Cuadrícula EIPOL.

4. El modelo de las dos caras de la evaluación

Este modelo se asocia al nombre de Robert E. Stake. Se llama así porque Stake habla de dos semblantes o caras de la evaluación: la descripción y el juicio.

Este modelo se relaciona directamente con la evaluación de los efectos en términos de los objetivos establecidos e implica completar dos matrices de datos como sigue:

	Matriz Descriptiva		Matriz de juicio	
	Intenciones	Observaciones	Estandares	Juicio
Antecedentes (Insumos)	X	X	X	X
Transacciones (procesos)	X	X	X	X
Resultados	X	X	X	X

La tarea del evaluador es encontrar datos para todas las celdillas en la tabla para comparar las observaciones con las condiciones y juzgarlas en términos de los estándares convenidos entre los organizadores del programa y los evaluadores. Se debe anotar que en el vocabulario de sistemas los antecedentes son insumos y las transacciones son procesos. En su puesta en marcha, el modelo ha usado muestras aleatorias estratificadas para recoger información específica en combinación con el método de estudio de casos.

Este modelo ha llamado la atención de los evaluadores respecto a la necesidad de definir estándares sobre los cuales puedan hacerse juicios, aunque el modelo mismo no ha resuelto el problema de la especificación de los estándares.

5. La evaluación responsiva

Con posterioridad, Stake se ha desplazado hacia el concepto de Evaluación Responsiva, un modelo de evaluación que se acerca más a las evaluaciones transaccional y naturalista. No es pre-ordenada (esto es, no está previamente definida por el evaluador, como especialista) sino que responde a las necesidades reales de las audiencias que requieren la información. No se focaliza en las intenciones del programa sino en las actividades reales. Su perspectiva es múltiple y utiliza las comunicaciones que naturalmente ocurren entre todos los implicados. No sólo busca recoger información sino captar el espíritu y el misterio del fenómeno en estudio. Por lo tanto es informal y reiterativa y enfatiza las descripciones fuertes. Como puede observarse es un enfoque muy humanista de la evaluación.

6. El modelo de la discrepancia

Este modelo fue propuesto por Malcom Provus, quien definió la evaluación como el arte de describir una discrepancia entre lo esperado y lo ejecutado en un programa.

Los principios básicos del modelo son: los estándares, la ejecución y la discrepancia. La tarea es comparar la ejecución con los estándares para determinar la discrepancia y así poder juzgar el valor o la adecuación de un objeto. El modelo sugiere también que busquemos las discrepancias en términos de cinco aspectos diferentes de un programa:

1. el diseño del programa
2. su instalación
3. el proceso de ejecución
4. el producto, y
5. el costo

Superficialmente el modelo parece algo racionalista, pero no lo es. En realidad, humaniza la evaluación y la hace responsiva por la manera como aplica los conceptos de estándares, ejecución y discrepancia. Por ejemplo, el evaluador ni establece estándares ni juzga las comparaciones entre estándares y ejecución, sino que solamente recoge los datos sobre la ejecución e indica las discrepancias. Es el cliente quien debe fijar los estándares aunque el evaluador le ayude a esclarecer la estructura del diseño del programa y, así, a establecer estándares apropiadas. El cliente deberá señalar que tipo de información le será más útil para tomar decisiones y debe hacer los juicios sobre las discrepancias.

Mientras que reconoce la utilidad del método experimental en algunos casos, el modelo muestra preferencias por los métodos descriptivos de la historia y la antropología y el método de estudio de casos de la sociología y la psiquiatría. Por su énfasis relativo en métodos naturalistas, sugiere que los evaluadores trabajen en equipos de modo que puedan confrontar las percepciones individuales y cuestionar los estándares que se aplican para describir las discrepancias.

El modelo pretende suministrar permanente información a los que toman las decisiones sobre la ejecución de un programa en marcha. También pretende proporcionar información que tiene una relación directa de uno-a-uno con las decisiones que realmente se toman. No obstante, los recursos requeridos para aplicar efectivamente este modelo son considerables en términos de personal, tiempo y dinero.

7. La evaluación transaccional

El modelo transaccional de evaluación tiene sus raíces en la psicología transaccional, la cual considera la percepción y el conocimiento como procesos de transacción. Estas transacciones ocurren entre individuos concretos en situaciones concretas; y el evaluador, como observador, es siempre parte del conjunto de transacciones. El modelo se asocia al nombre de Robert M. Rippey, quien ha desafiado a los educadores y capacitadores para que se concentren en los procesos educativos -el programa, la clase y la escuela- más que en las calificaciones obtenidas por los estudiantes.

Se centra en rendir cuentas sobre las responsabilidades -se pide a los agentes de cambio que se estudien a sí mismos, sus roles, los sistemas en los cuales desempeñan esos roles y los sistemas más amplios que constituyen el entorno de los sistemas en que se producen los cambios. Las metodologías recomendadas son informales. En palabras del mismo Rippey⁸

Una comparación con evaluaciones sumativas y formativas tradicionales muestra que el objetivo de la evaluación es diferente: el sujeto de la evaluación es el sistema, no el cliente o los servicios que el sistema ofrece. Las variables se relacionan con los aspectos sociales, psicológicos y comunicativos del sistema más que con los objetivos manifiestos. La información se devuelve en forma continua al sistema. El evaluador mismo es una parte más del sistema en funcionamiento. Las consideraciones convencionales de confiabilidad, validez y objetividad no son tan importantes como las de oportunidad, pertinencia y los efectos observables de generar información evaluativa. En primer lugar, la evaluación

intenta transformar la energía conflictiva del cambio en actividad productiva para clarificar los papeles de las personas involucradas en los cambios que produce el programa, no para generar nuevos conocimientos o para atribuir causalidad.

Deben hacerse notar los supuestos que se utilizan en el modelo transaccional de evaluación con respecto al paradigma básico y las posiciones de valor adicionales introducidas en el modelo. Es, en efecto, un modelo con una carga grande de valores. Enfatiza en la información sobre relaciones y nos exhorta a ser sensibles a las consecuencias no esperadas. También implica que la evaluación sea realizada colectivamente por los protagonistas y diseñadores de un programa de cambio y por las representantes de aquéllos a quienes probablemente éste afecte.

8. La evaluación libre de metas

La idea de una evaluación libre de metas fue introducida por Michael Scriven. El señaló que con nuestro énfasis sobre las metas establecidas, nuestra búsqueda se había centrado completamente en los efectos esperados -efectos que queríamos crear bajo las metas aceptadas del programa. Este enfoque llegó a ser tan excluyente que con frecuencia desarrollamos una visión como por entre un túnel: buscar la evidencia de los efectos esperados y no ver nada más.

El sugirió que debíamos buscar los efectos reales del programa, los que realmente hubieran ocurrido, ya fuera que hubiesen sido intencionales o no. Pensaba que esto podría hacerse si concebíamos una evaluación libre de metas, independientes de los objetivos establecidos para el programa. Entonces podrían combinarse los resultados de la evaluación centrada en objetivos y la evaluación libre de metas. El uso de una evaluación libre de metas no de sugerir una evaluación en el así llamado “estilo responsivo”. En efecto, Scriven es muy entusiasta de la evaluación sumativa y se interesa por las condiciones que los consumidores puedan utilizar. Tan sugiere que hagamos más evaluación de personal (esto es, de los maestros, trabajadores de campo, etc.) que las personas rindan cuentas de sus responsabilidades.

9. Enfoques investigativos a la evaluación

Jack D. Douglas⁹ ha analizado los métodos de los investigadores o los detectives para mostrar como las estrategias investigativas podrian usarse para exponer la verdad acerca de las personas situaciones sociales.

El modelo investigativo no supone la existencia de un mundo de cooperacion, apertura y veracidad sino de información errada, evasivas, mentiras enfrentamientos. Por eso, sugiere estrategias para captar una situación de evaluación, infiltrarse, construir relaciones amistosas y de confianza y usarlas en un proceso continuo de prueba y control.

El modelo *modus operandi* sugerido por, Michael Scriven es también método investigativo para estudiar relaciones de causa-efecto a través del examen secuencial. Este método reconstruye los procedimientos del historiador, el detective, el antropólogo y el ingeniero que apunta hacia los problemas. Este modelo se propone como sustituto de los enfoques experimentales y cuasi-experimentales cuando las situaciones de campo piden su utilización. En esencia, el método implica generar cadenas hipotéticas de eventos causa-efecto y eliminar aquellos que haya sido imposible que sucedieran. Este es, desde luego, el método característico del detective.

10. La evaluación iluminativa

Este modelo evolucionó como un claro rechazo al modelo de evaluación, “agrícola-botánico” enraizado en el paradigma científico. Se afirmaba que los grupos y comunidades no podían asignarse al azar a tratamientos como se hacía con las fincas y los terrenos agrícolas, y que a los seres humanos no se podía administrar tratamientos similares al de sembrar semillas en un terreno. En todo caso, los datos cuantitativos generados por el modelo agrícola-botánico sólo proporcionaban descripciones parciales del fenómeno.

Parlett y Hamilton¹⁰ construyeron este modelo iluminativo sobre la base de dos consideraciones importantes:

1. Los sistemas de instrucción, una vez adoptados, se convierten en sistemas vivos. Los sistemas vivos no se ajustan a las descripciones que se presentan en los catálogos. Ocurren modificaciones importantes en los programas entre como se dibujan en el taller a como se ponen en práctica realmente.
2. Los programas de capacitación y desarrollo no pueden separarse del ambiente de aprendizaje en el cual se realizan. Los actores en este ambiente y las estructuras del medio se constituyen en partes del sistema de instrucción.

Aunque retienen el uso de los métodos de muestreo y los cuestionarios y pruebas estructuradas. Parlett y Hamilton atrajeron nuestra atención hacia los métodos naturalistas centrados en la descripción y la interpretación. Las tres etapas en el proceso de evaluación sugieren incluir: (a) la observación de la situación educativa; (b) la selección de temas mediante un enfoque progresivo y una indagación intensiva; y (c) el análisis y la explicación.

11. La evaluación por expertos

El modelo de evaluación por expertos propuesto por Elliot W. Eisner¹¹ rompe claramente con el paradigma científico y rescata la tradición estética de las artes. Eisner dice que la enseñanza es un arte, y la escolarización un artefacto cultural. Entonces, ¿por qué no hacer evaluación al estilo de los entendidos? El afirma que, en efecto, sólo un experto que ha pasado toda una vida en el campo puede proporcionar evaluaciones mediante el uso sistemático de la sensibilidad perceptiva, la organización de la experiencia previa y la refinación de ideas perspicaces o intuiciones que serían imposibles de obtener de otra forma.

Eisner sugiere dos conceptos interrelacionados: (1) la experiencia en la educación (2) la crítica educativa para desempeñar las tareas de la evaluación educativa. La experiencia constituye el medio a través del cual pueda reorganizarse la forma del contexto y sus configuraciones interiores de modo que puedan tomarse decisiones inteligentes acerca del contexto. La crítica educativa es el arte de la revelación a través de la descripción, la interpretación y la evaluación.

La metodología que combina la experiencia y la crítica no es de manera alguna simple o romántica, y ciertamente, puede ser sistemática y rigurosa. Los críticos de la educación pueden aprender a buscar las cualidades permeables de la educación en las clases y en situaciones de capacitación; y pueden aprender a buscar los significados de indicios ocultos. Las cuestiones de confiabilidad y validez pueden manejarse por la corroboración estructural (validación mutua de una parte de los datos por el resto, siendo sustentada la totalidad por las partes que la constituyen); y por la adecuación referencial (la existencia de

una relación entre lo que dice el crítico educativo y el objeto de su crítica). También es posible hacer generalizaciones en el sentido de que la crítica en educación llevará a procesos más refinados de percepción en situaciones sucesivas; y suscitará nuevas anticipaciones en la mente del evaluador.

Los informes de críticas educativas se asemejan algo a los estudios de caso, pero los estudios de caso en forma de críticas educativas son diferentes en el sentido de que la crítica misma es una forma de arte. Como descubrimiento crítico, un informe educativo crítico crea una imagen vívida, que comunica a los lectores una comprensión de las realidades educativas.

12. El modelo de la evaluación como defensa

Este modelo es también llamado modelo de evaluación entre adversarios o modelo judicial. Como lo sugieren estas denominaciones el modelo utiliza procedimientos casi judiciales para conducir la evaluación. Es característico que se permita a dos grupos de personas, unos a favor y otros en contra de un programa, defender posiciones opuestas ante un jurado educativo respecto a los asuntos que van surgiendo y son seleccionados para el juicio. Se establecen las reglas y procedimientos probatorios y se permite el interrogatorio para comprobar lo declarado anteriormente. Es un juicio educativo con participación de un jurado.

Los proponentes del modelo citan varias ventajas del mismo: permite a los evaluadores desarrollar y utilizar procedimientos explícitos para generar y evaluar las estrategias de programas alternativos; proporciona un registro de las decisiones para luego establecer responsabilidades; no sólo acomoda los datos también sino también las percepciones, opiniones, sesgos y especulaciones, y puede implicar en el juicio a una gran variedad de interesados.

Por otra parte, hay quienes han encontrado serias fallas en el modelo, pues dicotomiza las posiciones en “pro” y en “contra” del programa. En la vida real, desde luego, no hay sólo dos sino muchos aspectos que considerar sobre un mismo asunto. El modelo convierte la evaluación en un evento competitivo. Como se decide de que lado estarán los grupos tirando al aire una moneda, con frecuencia los “abogados” no se asignan bien y ocurre que no ha convicción sobre las posiciones que defienden. Las habilidades de los jueces y los jurados son muy variables.

Debido a los muchos aspectos negativos del modelo y los altos costos que implica organizar el juicio, el formato que presenta “un caso en la corte” se ha transformado en “audiencias aclaratorias”. Se han eliminado los jurados, aunque todavía es usual emplear una especie de panel. Se puede llamar a testigos expertos que sirvan para sustentar posiciones en pro o en contra. Puede haber interrogatorios para comprobar declaraciones. Así se aclaran muchas cuestiones pero las decisiones sobre preferencias y modificaciones se deja que las hagan quienes escuchan a las partes.

13. El modelo de evaluación participativa

Frecuentemente se asocia el nombre de Paulo Freire, el educador brasilero y autor de *La Pedagogía del Oprimido* (New York: Herder and Herder, 1972) con la evaluación y la investigación participativas. En esta área los evaluadores diseminados por todo el mundo han realizado gran cantidad de trabajos durante los últimos diez años. Se han establecido redes de investigación participativa y se han llevado a cabo estudios evaluativos cuyos resultados han sido publicados.

La investigación o la evaluación participativa no es un esfuerzo científico de los profesionales sino un recuento existencial profundo de una experiencia en que toman parte todos los interesados conjuntamente, en colaboración. Los educandos se convierten en evaluadores y estos en estudiantes. Se decide en forma participativa acerca de las metas, los fines, los estándares y las herramientas de la evaluación. Cada participante contribuye con datos personales y recoge parte de los datos que deben obtenerse. Se emprende en forma colectiva el análisis de los datos. Los juicios también se emiten colectivamente.

En su discurso en el Instituto de Educación de Adultos de la Universidad de Dar es Salaam, en 1972, Paulo Freire presentó los posibles pasos para una metodología participativa:

1. El equipo de evaluación (o investigación) debe familiarizarse con investigaciones y evaluaciones realizadas previamente, sin importar los métodos utilizados en la mismas.
2. El equipo debe delimitar geográficamente el área de acción, aunque desde el punto de vista cultural no hay fronteras.
3. El equipo debe identificar las instituciones oficiales y organizaciones populares existentes en el área seleccionada y hablar con sus dirigentes.
4. El equipo debe decir honestamente a los líderes que ellos han venido a proponer la posibilidad de que todos los miembros de esa comunidad participen en las discusiones y trabajen conjuntamente.
5. Si los líderes aceptan, el equipo de evaluación debe celebrar reuniones no sólo con los líderes de las distintas instituciones sino también con otras personas que de alguna manera estén relacionadas con esas instituciones.
6. El equipo deberá discutir con la comunidad lo que se requiere para las reuniones, en las cuales pueden encontrarse unas treinta personas a diario o semanalmente para las discusiones. Estas reuniones pueden involucrar a casi todos los habitantes de la comunidad y prolongarse durante varias semanas. Lo importante es obtener la percepción que tiene la comunidad como una totalidad.
7. En esta etapa sociólogos, psicólogos, educadores y lingüistas deberán incorporarse al equipo de investigación o evaluación y visitar a cada grupo. Deberían registrarse las discusiones efectuadas en cada reunión. Debe instarse a que la gente hable si guardan silencio, de lo contrario, el papel del equipo debe ser sólo de asesoría. Uno de los miembros de la comunidad debe presidir estas reuniones.
8. Puede discutirse sobre muchos temas diversos como la justicia, la educación, el gobierno, la industria; pero todos referidos a las personas en el contexto de realidades concretas.
9. Cuando, los grupos más pequeños creen haber agotado los tópicos seleccionados para las discusiones cada grupo debe escribir sus conclusiones y presentarlas en sesión plenaria. Los relatores en dichas sesiones deben ser miembros de la comunidad, no los especialistas del equipo. Los trabajadores deben convertirse en intelectuales. Debe haber una discusión colectiva de cada uno de los informes de los grupos.
10. El equipo de evaluación debe hacer ahora un estudio crítico de las conversaciones de la gente. Este estudio debe ser interdisciplinario. Deben determinarse los diversos niveles en que la gente percibe la realidad y resolverse todas sus implicaciones. Esto no deben hacerlo los científicos sociales por su cuenta sino con los miembros de la comunidad.
11. Junto con la gente de la comunidad, el equipo de evaluación debe ahora hacer un borrador de una propuesta para emprender las acciones subsiguientes. El programa no debe resolverse para la comunidad sino con ella.

Con lo anterior debe quedar claro que la evaluación participativa no puede distinguirse del diagnóstico de necesidades o la concientización de la comunidad. También se diluye la distinción entre, evaluación e instrucción. La evaluación participativa proporciona a los participantes mayores oportunidades de elevar su nivel de conciencia y consolidar su sentido de poder y auto-valoración.

Como se señaló anteriormente, el Consejo Internacional para la Educación de Adultos ha trabajado mucho en la evaluación participativa y ha establecido una red de evaluación participativa a escala mundial. Sin embargo, no parece que haya habido desviaciones significativas de tipo metodológico. Los métodos han mantenido más o menos el espíritu de la listas de pasos propuestos por Paulo Freire y presentados aquí.

14. El modelo de evaluación (3-S)¹³ como estrategia para la situación específica

Antes de presentar nuestro modelo de evaluación 3-S, recordemos a los lectores que es inútil buscar el modelo de evaluación o una forma correcta de evaluar la alfabetización, la capacitación o el desarrollo. Como nos ha advertido Cronbach, un modelo puede ajustarse al “contexto del control” y al otro al “contexto de la acomodación”. Un agente alfabetizador puede estar usando más de uno de los modelos presentados aquí, en el contexto de un solo estudio evaluativo.

Otro punto importante que debe recordarse es que los modelos no se utilizan como fórmulas. Los modelos son para pensar. Raras veces le darán un conjunto de procedimientos cuyos pasos no puedan cambiarse. Si así lo hacen probablemente conducirán a cometer errores.

El modelo 3-S, el cual analizaremos a continuación, es un conjunto vacío que nos permite seleccionar el modelo apropiado o la combinación adecuada de modelos y enfoques que se vaya a usar en un programa o estudio evaluativo. La esencia de su concepción es la siguiente: no comience con un modelo de evaluación, sino con un problema. Analice el problema e identifique sub-problemas, piense acerca de cómo el problema o sus partes pueden desenvolverse con el transcurso del tiempo; y, finalmente, piense en el medio en el cual se desarrollará la evaluación. Las diferentes partes del problema de evaluación probablemente requerirán la utilización de distintos modelos y enfoques. Podrá ser necesarios hacer una encuesta y un estudio de caso en profundidad. Podrán requerirse pruebas de rendimiento de los educandos y análisis de contenido de documentos.

Las exigencias de tiempo pueden demandar tomar el pulso al programa mediante apreciaciones rápidas aún cuando idealmente, hubiese sido mejor hacer una evaluación más sistémica. Finalmente, el evaluador puede encontrarse en una situación en la cual no tenga calculadoras ni colegas que puedan ayudarle con el análisis de grandes cantidades de datos numéricos, donde no haya fotocopiadoras o mimeógrafos, o no se tenga papel para duplicar los instrumentos necesarios. Este modelo nos ayuda a pensar sobre las estrategias que habrá que seleccionar en situaciones especiales de la vida real sobre cómo hacer una “segunda mejor selección” cuando la mejor no es posible.

En otra publicación¹⁴ hemos enunciado los siguientes pasos en la puesta en práctica del modelo 3-S:

1. Articular las relaciones de fines y medios en el programa de cambio que se va a evaluar.
2. Generar perfiles de necesidades de información y asuntos relacionados con la evaluación.

3. Desarrollar una agenda de evaluación específica para la situación.
4. Elegir metodologías y técnicas realistas apropiadas.

El modelo 3-S ha influido en los enfoques de la planificación y la gestión de la evaluación que se presentan en este libro.

Cosas para hacer o para pensar

1. ¿Cuál de los dos paradigmas básicos analizados en este capítulo tiene más posibilidad de generar la información más útil en su programa de desarrollo, de alfabetización o de capacitación para el desarrollo? ¿O tendrá usted que usar una mezcla de ambos?
2. ¿Cuál o cuáles de los modelos descritos en este capítulo considera usted personalmente que le podría ser más útil actualmente? ¿Qué más le gustaría saber sobre este modelo para poder utilizarlo?
3. ¿Ha podido usted encontrar algún estudio evaluativo terminado que se ajuste del todo a alguno de los modelos descritos en este capítulo?

Notas de la lectura

Kuhn, T. S. *The structure of scientific revolutions* Chicago: University of Chicago Press, 1962.

En nuestra discusión en este capítulo el término “teoría” no ha sido definido ni explicado. Esto es así porque la literatura sobre evaluación habla con frecuencia de modelos de evaluación y pocas veces de teoría de la evaluación. Brevemente digamos que, en términos del status conceptual, la teoría cae entre el paradigma y el modelo. En su mejor sentido, la teoría es un conjunto de leyes y generalizaciones empíricas conectadas deductivamente. Un modelo es frecuentemente un diagrama esquemático que conecta la teoría con la práctica.

³ Aquí puede hacerse referencia al trabajo de Donald T. Campbell y Julian C. Stanley, *Experimental and quasi-experimental designs for researchers*. Chicago: Rand McNally, 1966.

⁴ Lincoln, Yvonna S y Guba, Egon G. *Naturalistic inquiry*. Beverly Hills, CA; Sage, 1985.

⁵ Cronbach, Lee J. et al *Towards reform of Jossey-Bass*. 1980.

H. S. Bhola. "Técnicas e instrumentos de evaluación". En: La evaluación de proyectos y campañas de alfabetización para el desarrollo. Instituto de la UNESCO para la educación. Antología UPN "Evaluación y seguimiento en la escuela LE' 94" pp. 134-139 Ed. Santiago de Chile (1992 pp. 186-198)

TEMA 2. Técnicas e instrumentos de evaluación

LECTURA: EL PROCESO EN UNA MIRADA: HERRAMIENTAS Y TECNICAS DE LA EVALUACION RACIONALISTA*

PRESENTACION

Es comprensible que las herramientas y técnicas de la evaluación racionalista (ER) sean sumamente racionalizadas. Los instrumentos de la ER son típicamente pre-estructurados y se someten a prueba previamente. Se desarrollan códigos detallados para las preguntas abiertas que se incluyan en los instrumentos. A los investigadores de campo se les advierte que sean impersonales con el fin de lograr objetividad. De antemano se establecen acuerdos sobre los métodos estadísticos y niveles de confianza que tendrán las inferencias que se hagan.

EL PROCESO DE UNA MIRADA:

HERRAMIENTAS Y TÉCNICAS DE LA EVALUACIÓN RACIONALISTA

Primero, será útil recordar en esta etapa que el proceso de planificación de la evaluación descrito antes, en el Capítulo 3 de la Parte II, se aplica para los tres enfoques analizados de generación de información: el SIG, la EN y la ER. Segundo, hay claros paralelos entre el diseño de un Sistema de Información para la Gestión y el de la ER. Los pasos peculiares en la progresión de la evaluación a la manera racionalista se muestra gráficamente en la Figura 8, que aparece en la página siguiente.

En este capítulo discutiremos cuestiones importantes acerca del diseño de instrumentos: la administración de estos en el campo; el cotejo (acumulación) de los datos recolectados, el procesamiento y presentación de los datos; y finalmente, el análisis de los mismos para someter a prueba las hipótesis o responder a las preguntas de la evaluación.

* H.S. Bhola, "Herramientas y Técnicas de la Evaluación Racionalista", en: La evaluación de proyectos, programas y campañas de alfabetización para el desarrollo. Instituto de la UNESCO. Editado en Santiago de Chile, 1992. Pp. 232-242.

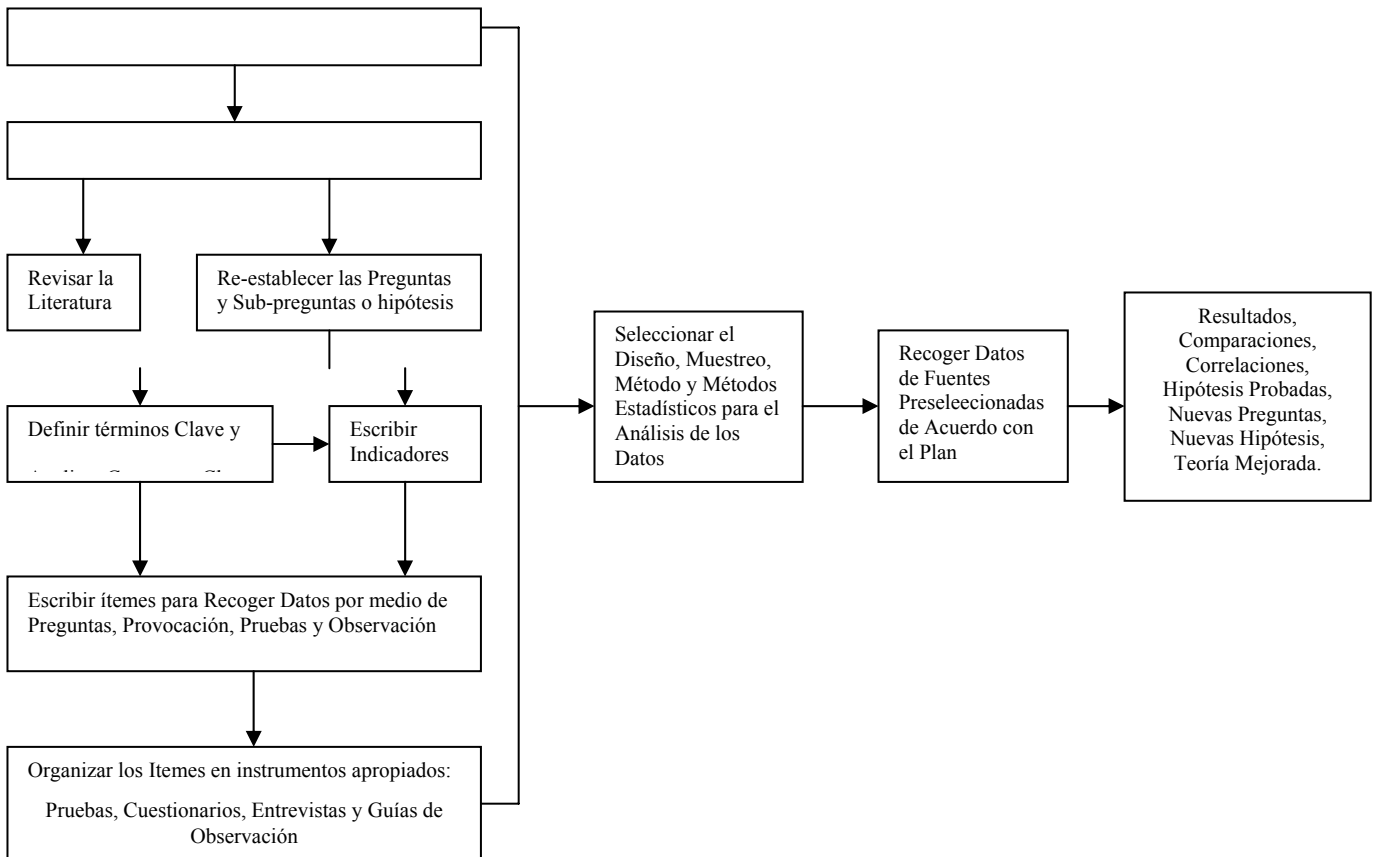


Figura 8. El Proceso de Evaluación

SECCION A: Herramientas e instrumentos

Cualquiera que sea la naturaleza del enfoque para la recolección de información con propósitos evaluativos, implica algunos tipos de recolección de datos. Habrá que ver, observar, preguntar, entrevistar, provocar reacciones y someter a prueba. Desde luego, hay diferencias esenciales en la forma como se recopila la información en la EN en comparación con la ER. Los instrumentos utilizados en la EN son inestructurados o muy poco estructurados. Hay muchas preguntas abiertas y los datos son cualitativos. Los instrumentos estructurados se consideran un mérito de la ER. Los datos son nominales u ordinales. Aún cuando se recogen “datos cualitativos”, estos se “cuantifican” para procesarlos como datos numéricos o nominales. Puede deducirse fácilmente que como el SIG típicamente guarda datos numéricos, tiene más afinidad superficial con la ER que con la EN.

Los cuestionarios estructurados

Como el nombre lo sugiere, estos son “estructurados” en cuanto a las preguntas que se formulan, las palabras exactas que deben usarse al presentarlas, la secuencia en que se preguntan y el formato en el cual las respuestas deben obtenerse y registrarse (debe

anotarse que los cuestionarios no son un simple conjunto de “cuestiones” con un signo de interrogación al final. Los cuestionarios incluyen escalas, ítems de elección múltiple y otros dispositivos para obtener y registrar las respuestas).

Los cuestionarios estructurados se distribuyen con frecuencia por correo. En casos especiales pueden hacerse a mano y, de hecho, pueden ser llenados también por el investigador de campo. Este será el caso cuando se usa, por ejemplo, un cuestionario con un grupo seleccionado de adultos analfabetos. En este caso, un cuestionario estructurado se convierte en una entrevista estructurada.

Los cuestionarios deben ser cortos y estar bien diseñados. Puesto que los informantes los llenarán independientemente, deben incluir instrucciones claras y fáciles de entender. Una corta introducción debe presentar el propósito del cuestionario y explicar como los datos que suministra el informante lo ayudara a el y a su comunidad en gneral. Debe asegurarse y garantizarse el anonimato del informante.

Los cuestionarios y las entrevistas son también en un sentido pruebas de conocimiento. La diferencia es que estas son pruebas del “conocimiento particular” que puede tener un individuo y lo quiera aportar. No es el conocimiento general de una materia sino conocimientos privados de una persona, información de la cual dispone la persona, sus percepciones, actitudes y opiniones de distinto tipo (como se ha indicado en otra parte, algunos ítems de las pruebas de rendimiento pueden a veces ocultarse en un cuestionario).

Adaptaciones locales de buenos cuestionarios

Generalmente los evaluadores tendrán que diseñar sus propios cuestionarios para que sean adecuados a los contextos sociales especiales y del programa en que se desarrolla el estudio evaluativo. Sin embargo, a veces pueden adaptarse cuestionarios sobre temas similares elaboradas por otros evaluadores en otros lugares. Muchos ítems útiles pueden tomarse de otros cuestionarios con pocas modificaciones.

La elaboración de buenos cuestionarios

Los buenos cuestionarios se elaboran en la perspectiva de objetivos claros. Preguntan lo que el evaluador necesita saber y evitan las preguntas innecesarias, pero no se olvidan las preguntas importantes. Siempre se solicita información demográfica estándar como sexo, edad, ocupación, ingresos, etc., de modo que se posibilite la interpretación de la totalidad de respuestas recibidas.

La elaboración de ítems para los cuestionarios presenta un conjunto adicional de problemas, ya que (1) se pregunta sobre conocimientos privados que los informantes pueden no querer compartir (2) se busca sacar opiniones y actitudes que es posible que los informantes no esten preparados para compartir honestamente. Las actitudes respecto a la planificación familiar, matrimonios entre personas de diferentes tribus, alimentos considerados tabú, no se expresan fácilmente con honestidad. El informante puede proporcionar respuestas “socialmente aceptables”. Puede decir a los evaluadores lo que supone que refleja la actitud apropiada en lugar de lo que realmente cree con respecto a algún aspecto particular de su mundo social o cultural.

Para esolver algunos de estos problemas, quienes elaboran los cuestionarios pueden intentar redactar ítems menos directos y preguntar los mismo de diferentes maneras en un mismo cuestionario.

Insistimos nuevamente en probar los cuestionarios antes de administrarlos a gran escala como parte del estudio evaluativo. La prueba previa señalará muchos problemas que puede tener el cuestionario.

La siguiente lista muestra algunos ejemplos de errores que cometen los evaluadores principiantes al redactar ítemes para cuestionarios. Muchos de esos problemas se captan en el proceso de pre-prueba del cuestionario. Con la práctica seguramente mejoraría la elaboración de ítemes para cuestionarios.

ITEMES

A un funcionario del distrito le preguntan: ¿Después que se ha comunicado la información a los jefes/asistentes en su área, cómo se trabaja al respecto?

Le preguntaron a un agente que trabaja a nivel local sobre nutrición: ¿A qué se dedica en las visitas al hogar?

Se le pregunta a un sujeto. ¿Cree usted que gozaba de buena salud durante el período que duró el curso de capacitación?

A un extensionista le preguntan en un curso de capacitación: ¿Fue útil su visita al agricultor?

Se solicita al director de una escuela que juzgue el compromiso con el trabajo entre el estudiante y el maestro, en términos de:

- Insatisfactorio
- por debajo del promedio
- promedio
- por encima del promedio
- sobresaliente

Comentarios

¿Puede el funcionario del distrito realmente responder? ¿No será mejor obtener esta información directamente de los jefes? ¿No estamos preguntando a informantes inadecuados?

¿No es está una pregunta demasiado general?

¿Entiende lo mismo por buena salud el sujeto y el evaluador? ¿Qué tal si el estudiante no estaba muy bien pero no tan enfermo como para perder clases por mucho tiempo? ¿No deberíamos preguntar en términos de los días que faltó debido a enfermedad?

¿Util para quién? ¿en qué forma? ¿Con base en qué clase de evidencia, usando qué criterios?

¿Cómo podemos garantizar que el evaluador y el director entiende lo mismo por compromiso con el trabajo? ¿Definimos el compromiso en términos de puntualidad o de

tener una sobrecarga de trabajo, o de ofrecer tutoría a los estudiantes con dificultades?
¿cómo adquirirá el director el conocimiento sobre el cual basar estos juicios?

El evaluador pregunta al extensionista local: ¿Son los materiales audio-visuales preparados localmente mejores que los preparados en otra parte?

Se pregunta a un asistente de cooperativas a nivel comunitario: ¿Cuántos de sus primeros estudiantes todavía practican las habilidades de lectura?

¿Significa “en otra parte” en otra localidad? ¿En la Oficina Central a nivel nacional?
¿Preparados por una firma comercial? ¿”Mejor” significa producido en otra parte? ¿Mejor en valores de producción o en términos de pertinencia de la instrucción?

¿No diría la mayoría de ellos que “muchos”?

¿No es éste una pregunta cargada?

Las entrevistas

Los evaluadores usan las entrevistas tanto en la indagación racionalista como en la naturalista. En el contexto del paradigma racionalista, las entrevistas son estructuradas o semi-estructuradas. Por semi-estructuradas entendemos entrevistas básicamente estructuradas que contienen algunas preguntas de sondeo que buscan mayores explicaciones.

Como indicamos antes, los cuestionarios estructurados se convierten en entrevistas estructuradas al ser aplicados en persona. Por lo tanto, la entrevista estructurada tiene los mismos problemas y suscita las mismas inquietudes respecto al diseño, la elaboración de ítemes y la presentación de los datos que el cuestionario estructurado. Pero, como las entrevistas se llevan a cabo en situaciones cara a cara., plantean otros problemas y desafíos adicionales. El entrevistado debe estar motivado para acceder a la entrevista e invertir el tiempo requerido para completarla. El entrevistador debe ser capaz de establecer confianza y empatía sin llegar a influir en las respuestas del entrevistado. En las localidades rurales de los países en desarrollo no siempre es posible apartar de sus ocupaciones al entrevistado (especialmente a las mujeres) para conversar un largo rato en privado. Por otra parte, el entrevistador debe asegurarse de que una entrevista individual con una joven madre no convierta en una entrevista familiar.

A veces lo que queremos son entrevistas familiares. Pero entonces debemos preparar y realizar una entrevista familiar. El punto es que una entrevista individual no se confunda con una entrevista familiar o grupal. También es posible que dos entrevistadores realicen una entrevista. Un líder de una comunidad rural puede ser entrevistado sobre su trabajo por un panel completo de entrevistadores.

Las escalas incluidas en los cuestionarios o las escalas independientes para registrar actitudes y opiniones

Como hemos indicado anteriormente, los cuestionarios no sólo incluyen preguntas. En un cuestionario puede aparecer una variedad de ítemes, incluyendo escalas. En su forma más sencilla una escala se presenta así:

SI	INDECISO	NO
ESTOY DE ACUERDO	NO ESTOY DE ACUERDO	ESTOY EN DESACUERDO

Una escala puede hacerse más sensible simplemente añadiendo más opciones ordinales, como en la siguiente:

ESTOY COMPLETAMENTE DE ACUERDO	ESTOY DE ACUERDO	NO ESTOY DE ACUERDO	ESTOY EN DESACUERDO O	ESTOY COMPLETAMENTE EN DESACUERDO
--------------------------------	------------------	---------------------	-----------------------	-----------------------------------

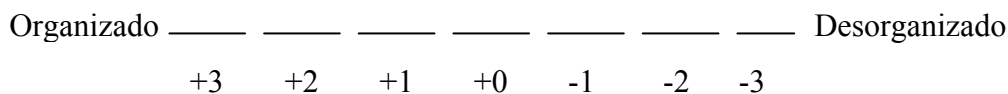
A las opciones ordinarias intermedias de la escala no siempre hay que darles definiciones y de hecho, pueden dejarse en blanco. Fíjese en la siguiente escala que usa siete opciones ordinales sin rótulos con dos extremos polares opuestos:



Estas escalas pueden convertirse en escalas multidimensionales utilizando muchas dimensiones bipolares, tales como las siguientes:

creativo-no creativo	organizado-desorganizado
difícil-fácil	relevante-irrelevante
flexible-inflexible	práctico-no práctico
emocionante-aburrido	activo-pasivo
fuerte-débil	exigente-no exigente
científico-artístico	envolvente-alienante
objetivo-subjetivo	modificable-inmodificable
	motivante-alienante

Estas escalas pueden ser analizadas en conjunto para obtener una visión más precisa de la estructura referente a la actitud o de los valores de un individuo. A veces puede darse valores numéricos a estas escalas como en la siguiente:



Esto permite la cuantificación de los datos recogidos utilizando rótulos cualitativos. Al ser aplicadas a un grupo, estas escalas pueden utilizarse para describir las estructuras de los grupos, al obtener los porcentajes de las respuestas.

Por ejemplo:

Creativo	3.7	22.2	26.6	23.8	18.3	5.8	No creativo
	%	%	%	%	%	%	

Observaciones de campo

La observación de campo, repetimos, es una estrategia de recolección de datos que puede utilizarse al seguir cualquiera de los dos paradigmas, el racionalista y el naturalista. En el paradigma racionalista la observación de campo puede basarse en el muestreo aleatorio y puede ser altamente estructurada. En la tradición naturalista, la observación de campo es inestructurada y propositiva. Pueden hacerse observaciones participantes o no participantes. Los evaluadores desean hacer observaciones de campo para sentir directamente la realidad sin que un intermediario tenga que ver e interpretarla por ellos. Sin embargo, la observación no es simplemente cuestión de abrir nuestros ojos y oídos a la gente en las situaciones de la vida real. Tenemos que entrenar nuestros ojos y oídos y aprender a registrar nuestras observaciones. Los diarios, listas de cotejo, los mapas y diagramas, guías, escalas sociométricas, escalas de calificaciones y cámaras pueden utilizarse para registrar observaciones.

Las guías de observación no son fáciles de elaborar y pueden deslizarse en ellas gran cantidad de errores. Examinemos los siguientes ejemplos:

ITEMES

COMENTARIOS

¿Pueden los alumnos sentados en la parte posterior de la clase escuchar al maestro-estudiante?	¿Puede observarse esto? ¿O tenemos que preguntar a los alumnos sentados en los asientos de atrás si oyen? ¿O debería el evaluador tratar de escuchar desde la parte trasera del salón?
¿Habla el maestro-estudiante con confianza en sí mismo?	¿Qué debemos mirar para observar manifestaciones de confianza en sí mismo?
¿Qué situación económica tienen los que han solicitado préstamos?	¿Podemos “observar” la situación económica como tal?
¿Cómo utilizaron los fondos quienes obtuvieron préstamos de la sociedad cooperativa; tomaron otra esposa, pagaron pensiones de los niños, comenzaron a beber más o compraron ropa nueva?	¿Cómo podemos observar esta historia de comportamientos durante una visita o un corto período de observación? Este tipo de información tendrá que ser recogida por otros medios alternativos.
Una guía de observación busca observar:	¿Es posible observar esto?
▪ actitudes de las personas antes de que comiencen una reunión pública; y	¿Cambiar las actitudes durante una reunión pública?
▪ actitudes de las personas durante y	¿Se muestran las actitudes en las caras de las

después de la reunión	personas?
¿Lleva la sociedad cooperativa los libros exigidos por la ley?	Está bien: pero ¿no es esto una cuestión de auditoría más que de observación?

Registros y documentos

Los registros y documentos son fuentes de datos importantes para el evaluador. El análisis de los registros y documentos puede ser cuantitativo (adecuado el paradigma racionalista) o cualitativo (adecuado el paradigma naturalista).

La ética de la compra de datos

Con frecuencia surge esta interrogante: ¿Debe un evaluador pagar a los informantes en un estudio evaluativo? No hay una simple respuesta de “Sí” o “No”. La producción de conocimiento es una función social; y en el caso de un estudio de evaluación, la utilización social de la información evaluativa puede ser bastante clara tanto para los evaluadores como para los informantes. Si el evaluador está trabajando a nombre del gobierno o de una agencia voluntaria sin ánimo de lucro, la evaluación está sirviendo un interés público. Los informantes como buenos ciudadanos comprometidos, deben participar gratis en el estudio evaluativo.

Sin embargo, si se coloca a un sujeto en una situación de tener que elegir entre trabajar en la construcción o participar en el estudio evaluativo, se le debe pagar para compensarlo por la pérdida del jornal.

Cosas para hacer o para pensar

1. Desarrolle una lista detallada de afirmaciones verdaderas, principios, destrezas y actitudes que usted quisiera que hubieran aprendido los que participan en el curso de capacitación hacia el final del mismo.
2. ¿Le han entrevistado recientemente como parte de una evaluación o algún tipo de encuesta? ¿Qué recuerda que haya sido bueno en la entrevista? ¿Qué le pareció irritante o inaceptable? ¿Fue capaz el entrevistador de ganar su confianza?
3. Elabore una guía de observaciones sobre “Los Hábitos de Trabajo en la Oficina”. Ensáyela observando a un colega. Pídale a su colega que la use para observarle a usted.

H. S. Bholá. “Herramientas y técnicas de la Evaluación Naturalista”. En: La evaluación de proyectos y campañas de alfabetización para el desarrollo. Instituto de la UNESCO para la educación. Antología UPN “Evaluación y seguimiento en la escuela LE’94” pp. 164-172. Ed. Santiago de Chile (1992 pp. 186-198).

LECTURA: EL PROCESO EN UNA MIRADA: HERRAMIENTAS Y TÉCNICAS DE LA EVALUACIÓN NATURALISTA*

PRESENTACIÓN

El hecho más importante de la evaluación es que no presenta excusas por el uso del individuo humano como instrumento de recolección de datos. No tiene pretensiones de construir instrumentos “confiables” que recojan datos “objetivos”. En efecto, los instrumentos más ampliamente utilizados en la EN son la entrevista inestructurada y la observación participante. Los materiales documentales también son muy usados. El evaluador naturalista no pretende permanecer por fuera de su estudio evaluativo. Con frecuencia se hace uso de un diario reflexivo para registrar las ideas y experiencias personales.

EL PROCESO EN UNA MIRADA:

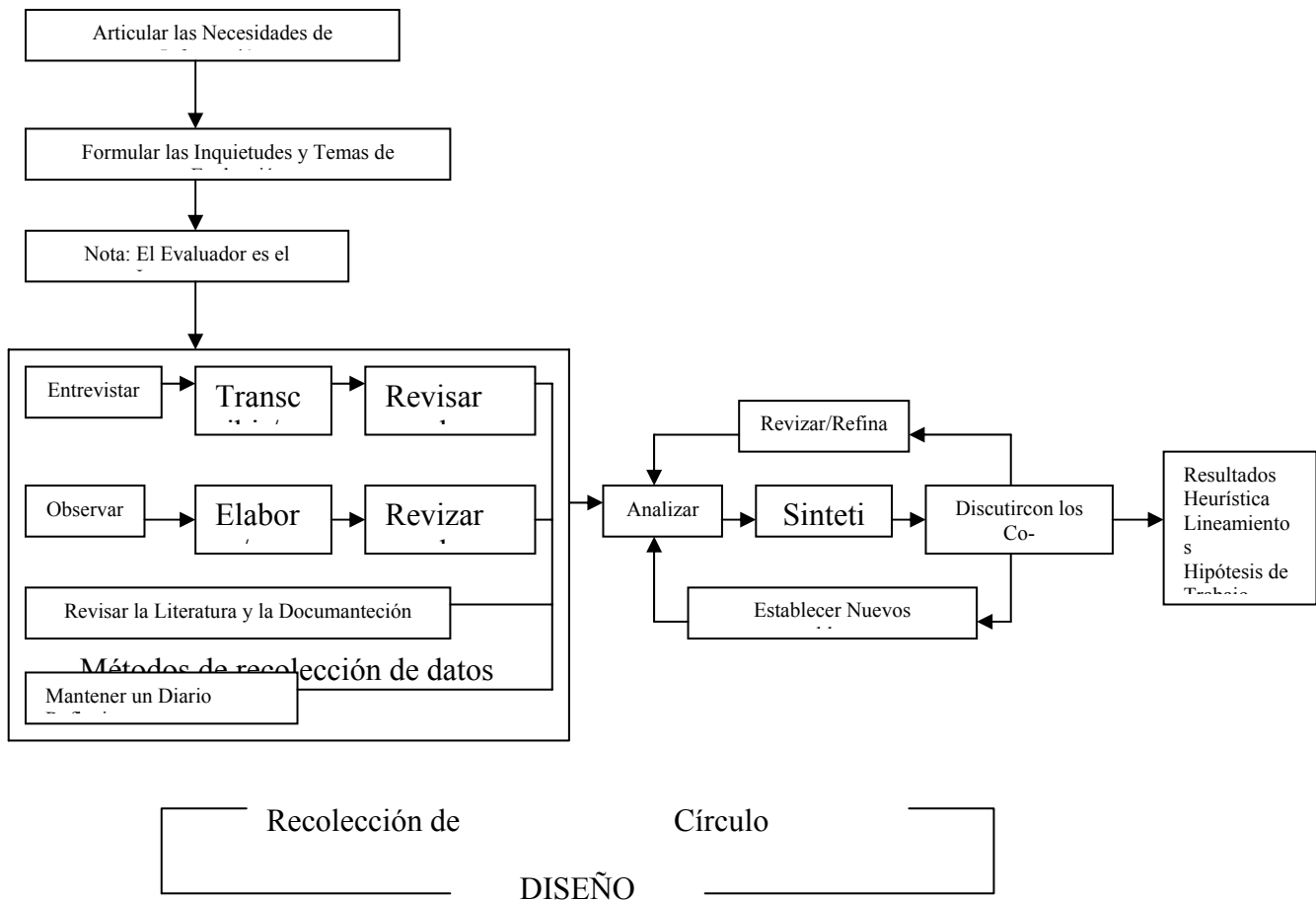
HERRAMIENTAS Y TÉCNICAS DE LA EVALUACIÓN NATURALISTA

La evaluación naturalista busca estudiar la vida como una totalidad, con todas sus complejidades, tal como esta enraizada en su contexto y es vivenciada por aquellos que están inmersos en la realidad que se estudia. Es un encuentro personal, no cualquier examen desapasionado.¹ Por lo tanto, no debe sorprendernos que el evaluador naturalista quiera ver, preguntar e interactuar para reunir todas las cosas.

El instrumento básico en la EN es el humano -el evaluador mismo. Puesto que la EN esta interesada en los significados del actor y su contexto, casi siempre se utiliza la entrevista y la observación inestructuradas de modo que puedan hacerse ricas descripciones (que dan al lector el sentimiento de haber estado ahí). Estas descripciones deben tener resonancia y ser coherentes. Otras herramientas y técnicas son las del análisis de contenido de los documentos.²

La Figura 1 de la página siguiente, presenta en forma gráfica la progresión característica de la indagación naturalista. Como puede observarse en, el diagrama, la fase de recolección de datos típicamente comprenderá entrevistas, observaciones, análisis documentales y el uso de diarios reflexivos. Como parte del análisis de los datos; estos datos serán sometidos al sub-diseño del círculo hermenéutico para obtener los resultados requeridos, en términos del estudio y/o la heurística, lineamientos guía, hipótesis de trabajo o modelos.

* H. S. Bhola, “Herramientas y Técnicas de la Evaluación Naturalista”, en: La evaluación de proyectos, programas y campañas de alfabetización para el desarrollo. Instituto de la UNESCO para la educación. Editado en Santiago de Chile, 1992. pp. 186-198.



A continuación incluimos algunas notas sobre algunas herramientas típicas de la EN.

Figura 1: El Proceso de Evaluación
La entrevista en la EN

La entrevista en la EN es inestructurada en el sentido que no pregunta al entrevistado una serie de preguntas estándar ya hechas pero es estructurada en un sentido más profundo y en un nivel superior. En términos de Paulo Freire, tal entrevista puede describirse como dialógica. Su propósito no es obtener respuestas a las preguntas, sino permitir al entrevistado describir su mundo en sus propias palabras. Sugiriendo cuestionamientos genuinos, haciendo comentarios serios y tranquilizadores y proporcionando refuerzos, el entrevistador puede llegar a interactuar profundamente con el entrevistado quien podrá entonces formular coherentemente lo que para él significan las realidades de su entorno.³

El entrevistador comienza por plantear temas generales y exponer las raíces problemáticas de ciertos temas. Después de establecer los límites generales del tema de la entrevista, el entrevistador naturaliza deja que el entrevistado se haga cargo del curso de la misma. El entrevistador escucha con interés y simpatía, animando al entrevistado a profundizar sus explicaciones, a volver a algún punto anterior, a seleccionar, juzgar y tomar posiciones. Una entrevista como ésta dura normalmente varias horas. La evaluación naturalista es, efectivamente, un trabajo intensivo.

Las entrevistas en profundidad pueden ser registradas por el entrevistador tomando notas en presencia del entrevistado o inmediatamente después, cuando el entrevistador tiene todo fresco en su mente. Si el entrevistado lo permite sería preferible grabar la entrevista.

Entrevistas grupales focalizadas

El evaluador naturalista puede con frecuencia hacer uso de entrevistas grupales focalizadas en temas particulares. Esta técnica se ha utilizado en la investigación de mercados en los Estados Unidos por más de veinte años. Su objetivo es recoger información profunda a través de la discusión grupal, llegando así a conocer los pensamientos, percepciones, sentimientos y actitudes de personas conocedoras de un programa.

El grupo focal es, característicamente, un grupo homogéneo de siete o más personas en interacción “dirigida” unos con otros. La dirección del grupo está en manos de un moderador. Un observador está también presente. La discusión comienza con una lista de inquietudes, llamadas “preguntas estímulo”. Generalmente se graban estas entrevistas, que normalmente toman de 45 a 60 minutos.

Como en todas las técnicas de dinámica de grupos, el modelador tiene que dirigir al grupo de modo que se produzcan bastantes ideas, que estas sean variadas y de buena calidad, que los participantes no se aparten del tema, que se mantenga la espontaneidad y se compartan las ideas cuando estas no estén del todo desarrolladas; que los participantes produzcan ideas y no simplemente reaccionen a las ideas de otros, y que la personalidad y el estatus de los participantes no influyan en el grupo.⁴

La observación de campo

La observación de campo es una estrategia de recolección de datos que pueden usarse tanto para las evaluaciones racionalistas como para las naturalistas.⁵ En la tradición de la EN, la observación será inestructurada y se hará sin prisa. Puede ser una observación participante o no-participante.

En la observación deben separarse las fases de registro y de interpretación. El observador debe registrar lo que ha visto. Lo que pudo significar debe reservarse para la interpretación posterior. Sea que se utilicen en la EN o en la Er, pueden elaborarse instrumentos sistemáticos en cada caso para registrar los datos de la observación.

La realización de estudios de rastreo y cronologos

Un estudio de rastreo, como lo indica su nombre, sigue el rastro del progreso de una persona a través del tiempo. Así, un estudio de rastreo de un neo-letrado parecerá una corta biografía de esta persona, registrando lo que la alfabetización puede haber hecho con la vida del neo-letrado, como aplica sus habilidades alfabetas en el trabajo y la vida en general.

El profesor Myrtle Scott, colega de este autor en la Universidad de Indiana, ha desarrollado el concepto de cronologo, el cual implica el seguimiento de los sujetos durante un período de tiempo determinado -una hora, un día, una semana- en su habitat natural (una comunidad, una escuela, un hospital) y el registro de lo que hacen. Estos cronologos son de gran utilidad para el estudio de los papeles y funciones emergentes de los agentes del desarrollo y de la alfabetización a nivel de base.

Análisis de registros y documentos

El análisis de registros y documentos puede realizarse tanto en la ER como la EN.⁶ En la primera, el análisis de contenido implica la selección aleatoria del contenido y el uso de técnicas estadísticas para llegar a aseveraciones generales. En la EN el análisis de contenido está más interesado en la agregación de significados y la cristalización de temas que puedan estar insertos en diversos documentos.

Medidas no reactivas⁷

Este tipo de medidas también puede usarse en estudios evaluativos ER y EN. Como sugiere su nombre, estas medidas no influyen en los informantes. Se observa, por ejemplo las huellas de comportamiento que han dejado. El estado de la cartilla que usa el educando puede ser una indicación de cuanto la ha leído. La forma de vestirse el trabajador de campo puede indicar la “distancia social” que mantiene con respecto a las personas a las cuales busca servir. La basura que bota una familia puede decirnos mucho sobre sus hábitos de compra y consumo -¡sin que se den cuenta!

El diario reflexivo

El diario reflexivo es exactamente lo que sugiere su nombre -un diario en el cual se registran las reflexiones (Lincoln y Guba usan este término para indicar que el evaluador esta registrando en el diario gran cantidad de información acerca de sí mismo como instrumento humano de evaluación. El evaluador mantiene un diario en el cual registra preferiblemente a diario, (i) el programa de cada día y los aspectos logísticos del estudio; (ii) anota las decisiones metodológicas que va tomando día a día y por que fueron necesarias; y (iii) sus reflexiones personales sobre las experiencias y las expectativas, conflictos de valores y colaboraciones con las partes interesadas, el aburrimiento que le produce el trabajo y la emoción de las iniciativas que surgen. El diario reflexivo se convierte así en el super-ego del evaluador, una ayuda a la memoria y una fuente de control de sus interpretaciones.

Casos de métodos de EN en acción

La teoría y la metodología de la EN están todavía emergiendo en un proceso de discusión. Se combina una variedad de métodos como los presentados arriba, manteniendo los supuestos subyacentes de la indagación naturalista (esto es, la realidad es una construcción social) y su objetivo no es formular proposiciones universales a manera de leyes, sino “aseveraciones fundamentadas” en contextos particulares.

Hay pocas evaluaciones naturalistas disponibles en el campo de la alfabetización que pudieran presentarse como modelos a seguir. Sería útil, no obstante, esbozar los enfoques metodológicos y los métodos utilizados por algunos evaluadores en el campo educativo, quienes afirman que realizan análisis naturalistas o que tienen una orientación naturalista. Por consiguiente se han esbozado dos casos.

Los lectores serán capaces de derivar estos ejemplos algunas ideas iniciales sobre lo que significa hacer un estudio evaluativo con orientación naturalista.

CASO 1

John C. Creswell y sus seis asociados⁸ hicieron un análisis naturalista del papel relacionada con el desarrollo de la facultad que desempeñan los decanos de departamento en situaciones de educación superior. Estos investigadores no pueden ser considerados “puristas” de la metodología de la EN pero definieron su trabajo como “un proceso de investigación con una fuerte orientación naturalista, la cual aísla inductivamente el problema estudiado, descubre un marco teórico a priori en lugar de verificarlo, y reporta los resultados hasta la fecha para ser sometidos a una posterior exploración cualitativa y a una investigación cuantitativa”.

El método de investigación del análisis naturalista operacionalizado para este estudio comprendió cuatro fases:

Fase 1: Establecimiento de la dirección para la investigación.

Fase 2: Caracterización de la muestra y desarrollo de procedimientos de entrevistas para el estudio nacional.

Fase 3: Realización de las entrevistas telefónicas.

Fase 4: Análisis de los datos de los informes verbales de los casos entrevistados.

Las acciones detalladas de cada una de las fases enumeradas fueron:

Fase 1

Identificación de las dimensiones para estudiar el papel del decano. Esto incluyó algunas preguntas generales como: ¿Cuáles son las características demográficas de los decanos “efectivos” y se sienten ellos responsables por contribuir al crecimiento y desarrollo de la facultad? ¿Qué clase de situaciones de la facultad requieren su apoyo y que enfoques utilizan ellos? ¿Cómo es la ayuda del decano moldeada por variables contextuales del ambiente de trabajo académico tal como la etapa de la carrera del miembro de la facultad, si la asistencia es iniciada por el decano o miembro de la facultad: diferencias por disciplina y diferencias institucionales?

Uso de 99 actividades que los decanos realizaban.

Desarrollo de un documento conceptual

Reacción al documento por parte de líderes nacionales reconocidos en el campo del desarrollo de la facultad.

Realización de un estudio piloto, el cual dio oportunidad para someter a prueba la técnica de la entrevista telefónica.

Dirección de dos disertaciones doctorales.

En su conjunto estas acciones tuvieron como resultado el surgimiento de áreas de cuestionamiento tales como características de los antecedentes de los decanos; situaciones o

asuntos de la facultad; prácticas de desarrollo de la facultad, en general, y condiciones de los departamentos.

Fase 2

Búsqueda de representatividad en la muestra.

Desarrollo de una guía de entrevista semiestructurada; y entrevistadores capacitados para lograr una administración consistente y precisa de la guía.

Uso de formas de codificación para registrar las respuestas a las entrevistas.

Soicitud a las personas de que “¡identificarán de 3 a 5 decanos de facultades que sobresalieron en ayudar a la facultad en su crecimiento y desarrollo” en el primer ciclo de entrevistas.

Fase 3

Realización y registro de entrevistas durante un período de más de nueve meses, cada una de las cuales duraba un promedio de 45 minutos.

Control de confiabilidad entre entrevistadores.

Fase 4

Iniciación del análisis tan pronto como hubo suficientes casos disponibles. Sucedió que:

Las primeras 50 entrevistas se utilizaron para desarrollar los procesos de análisis.

Las preguntas abiertas de estos 50 casos se “rigieron” en categorías preliminares y se clasificaron de acuerdo al enfoque de la pregunta.

Entonces se incorporaron al análisis los 135 casos restantes.

Las “historias” variadas ayudaron a identificar tipos de “situaciones de la facultad” y de “enfoque de los decanos”.

Triangulación entre el equipo de investigación para recategorizar las situaciones y/o enfoques al examinar los casos más extremos, maximizar similitudes y diferencias y desarrollar prototipos.

Resultados: una matriz conceptual.

Todos los datos utilizados eran de auto-informes.

CASO II

En una reciente evaluación de los programas de la Organización de Alfabetización de Adultos de Zimbabwe (ALOZ) realizada por Bhola y Muyoba, la metodología tenía una orientación naturalista. Los procesos y los pasos se describen a continuación:

El contrato de evaluación había enumerado preguntas generales que necesitaban respuesta: ¿Cómo había afectado las vidas de aquellos a quienes querían servir los programas de alfabetización de adultos y de generación de ingresos de ALOZ? ¿De qué manera el trabajo de alfabetización había contribuido a la promoción de los esfuerzos de alfabetización a nivel nacional en Zimbabwe? ¿Cómo contribuyó la asistencia proporcionada por USAID/ZIMBABWE a ALOZ durante los primeros años de la década de los 80 a cumplir la misión de ALOZ?

Se realizó un análisis detallado de la documentación relacionada con la estrategia de Zimbabwe y del papel asignado a la alfabetización en la estrategia; la política nacional de alfabetización y la ejecución de la campaña nacional de alfabetización del gobierno, la misión de ALOZ en relación con la promoción de la alfabetización y su materialización: y las metas y los objetivos de la subvención de USAID/ZIMBABWE a ALOZ.

Se diseñó un juego de 12 cuadros para desarrollar un retrato numérico del trabajo y los logros de ALOZ durante el período que duró la subvención de USAID.

Se emprendió un extenso trabajo de campo para desarrollar una descripción de cómo los proveedores de los programas de ALOZ y los participantes y otros beneficiarios habían vivenciado el programa.

Se cubrió una muestra intencional de las localidades típicas para abarcar diferentes lenguas (Shona y Sindebele) y diferentes realidades socio-económicas (área urbano-industriales, plantas manufactureras, agro-industrias, zonas rurales, asentamientos rurales, etc.).

En cada localidad, se entrevistó a toda la gama de personas involucradas, entre ellos los participantes, sus esposas y otros miembros de la familia, beneficiarios indirectos de las actividades de generación de ingresos en las comunidades, maestros, supervisores, capacitadores, empleados de agencias auspiciadoras y colaboradoras, etc. Se utilizaron guías de entrevista semiestructuradas como base para las conversaciones con los informantes.

Primero, se elaboraron estudios de caso completos -uno por cada una de las localidades incluidas, usando las respuestas de categorías pertinentes de personas involucradas. Luego se usaron todos los datos -los datos numéricos organizados en cuadros y los estudios de caso- de varias formas para responder las preguntas específicas formuladas por los que toman las decisiones.

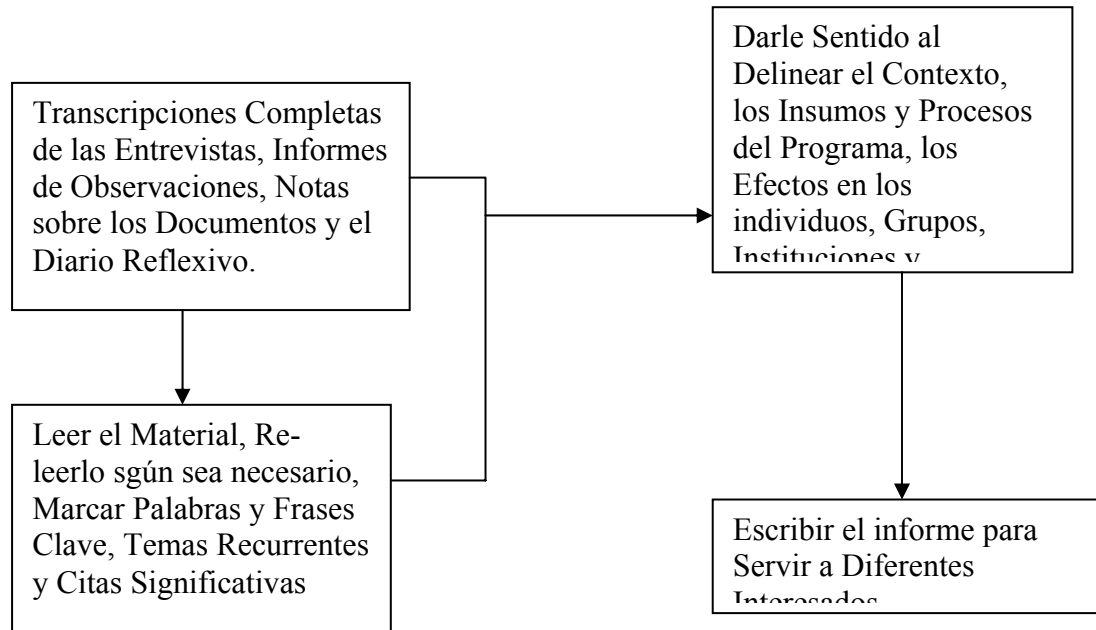


Figura 2: Focalización de la EN en el Procesamiento y

El Análisis de los datos en la EN

En análisis de los datos en la EN es de cierta manera un proceso mucho más desafiante que el análisis estadístico. No se trata solamente de una agregación de números, sino en la generación de significados y la búsqueda de los patrones más abarcadores en los cuales residen esos significados.

El primer paso en el análisis de los datos de EN es una inmersión total en los datos que ya se han recogido. El evaluador debe leer y re-leer las transcripciones de entrevistas, informes de observaciones, notas del análisis documental y el diario reflexivo, (ver la Figura 2 en la página anterior). Las palabras y frases clave, así como los temas recurrentes deben anotarse en tarjetas y también las citas importantes de comentarios hechos por las personas involucradas. Mediante un proceso de síntesis, usando el modelo CLER y el formato “antes-después” deben reportarse los cambios en las vidas de los individuos, grupos, instituciones y comunidades. Desde luego, no hay fórmulas estándar en el análisis de los datos de las EN, pero los evaluadores que usan esta modalidad están seguros de sacar provecho de la experiencia al tratar de darle sentido al mundo con base en los datos recogidos.

Cosas para hacer o para pensar

1. Use una grabadora para grabar su entrevista con un campesino para averiguar por qué él no puede seguir todos los consejos que le da el agente extensionista. Escuche de nuevo la entrevista. ¿Quién habla más, usted o el campesino? ¿Son sus preguntas un poco impacientes? ¿Está usted realmente escuchando?

2. Suponga que usted está interesado en aprender sobre el nivel de salud de una comunidad. ¿Qué observaría usted? Compare sus observaciones con las que haga un promotor de salud.

Notas de la lectura

¹Para una discusión general sobre la investigación de campo ver Johnson, J. M., *Doing field research*. New York, NY: The Free Press, 1975. Puede encontrarse una discusión de la teoría fundamentada y los avances en el concepto de ésta en Glaser, Barney y Strauss, Anselm L., *the discovery of grounded Theory*. Chicago, IL: Aldine Publishers, 1967; y en Glaser, Barney, *Theoretical Sensitivity: Advances in the methodology of grounded theory*, Mill Valley, CA: sociology Press, 1978.

² Algunas referencias son:

Bogdan R, y Biklen, S.K. *Qualitative research for education*. Boston, MA: Allyn and Bacon, 1982. Patton, Michael Quinn. *Qualitative evaluation methods*. Beverly Hills, CA: Sage, 1980. Miles, Matthew y Huberman, A. Michael. *Qualitative data analysis: A source book of new methods*, Beverly Hills, CA: Sage. 1984.

³ Gorden, R., *Interviewing*. 3rd ed. Homewood, IL: Dorsey Press, 1980 y Dexter, L.A., *Elite and specialized interviewing*. Evanston, IL: Northwestern University Press, 1970, son algunas referencias útiles para hacer entrevistas. También debe uno referirse a la literatura sobre las entrevistas etnográficas. La etnografía utiliza la descripción como componente importante de la recolección de datos. El énfasis está en la perspectiva émica (del interior) para poder comprender la realidad de los sistemas sociales: y al mismo tiempo proporciona el contexto de modo que se entiendan los datos apropiadamente. Ver Fetterman, David M., *Ethnography in educational evaluation*. Beverly Hills. CA: Sage, 1984; Spindler, George y Louise *Interpretative ethnography of education at home and abroad*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc., 1987; y Spradley, James P., *The ethnographic interview*. New York, NY: Holt, Rinehart and Winston, 1979.

⁴ krueger, Richaed A. *Focus group interviewing: step by step intructions for extension workers*. Minnesota Agricultural Extension Service. (320C Vocational Technical Building, 1954, Buford Avenue, St. Paul, Minnesota, 55108), 1985. También *Qualitative Research Council of the Advertising Research Foundation. Focus groups: Issues and Approaches*. New York, NY: Advertising Research Foundation, inc., 1985.

⁵ McCall, G.J. y Simrnons, J. L. (eds.) *Issues in parncipant observation: A text and reader*. Reading, MA: Addison-Wesley, 1969. También Spradley, J. P., *Participant observation*. New York, N.Y: Holt, Rinehart y Winston, 1980.

⁶ Krippendorff, Klaus. *Content Analysis*. Beverly Hills, CA: Sage, 1980.

⁷ Ver Webb, E.J. et al. *Unobtrusive measures*. Snokie, IL: Rand McNally, 1966 para una discusión sobre estas mediciones. Una discusión más reciente se presenta en Sechrest, Lee (ed.) *Unobtrusive measures today*. *New Directions for Methodology of Behavioral Sciences*, No. 1, San Francisco, CA: Jossey-Bass, 1979.

⁸ Creswell, John W, et al. “The faculty development role of department chairs: A naturalistic analysis”. Investigación presentada en el Annual Meeting of the Association for the Study of Higher Education, Baltimore, Maryland, noviembre 21-24, 1987.

⁹ Bholá, HS. y Muyoba, G.N. The role of the Adult Literacy Organization of Zimbabwe (A) in promoting Universal Literacy – A Retrospect and a Prospect. Harare. ALOZ/USAID/ZIMBABWE, 1989. Una descripción interesante de la ejecución del enfoque naturalista se encuentra en Valbuena Paz, Antonio y González Olivares, Guido. “Case Study of CESAP Programme: “Mucuchies Peasant Programme”, Hamburg: UNESCO Institute for Education (UE) project PRG 5. 14/4.53, Document 10, junio 1990. Mimeo.