



Mi Universidad

LIBRO

EVALUACIÓN DEL APRENDIZAJE

DOCTORADO EN EDUCACIÓN

TERCER CUATRIMESTRE

Marzo-Abril

Marco Estratégico de Referencia

Antecedentes históricos

Nuestra Universidad tiene sus antecedentes de formación en el año de 1979 con el inicio de actividades de la normal de educadoras “Edgar Robledo Santiago”, que en su momento marcó un nuevo rumbo para la educación de Comitán y del estado de Chiapas. Nuestra escuela fue fundada por el Profesor Manuel Albores Salazar con la idea de traer educación a Comitán, ya que esto representaba una forma de apoyar a muchas familias de la región para que siguieran estudiando.

En el año 1984 inicia actividades el CBTiS Moctezuma Ilhuicamina, que fue el primer bachillerato tecnológico particular del estado de Chiapas, manteniendo con esto la visión en grande de traer educación a nuestro municipio, esta institución fue creada para que la gente que trabajaba por la mañana tuviera la opción de estudiar por las tardes.

La Maestra Martha Ruth Alcázar Mellanes es la madre de los tres integrantes de la familia Albores Alcázar que se fueron integrando poco a poco a la escuela formada por su padre, el Profesor Manuel Albores Salazar; Víctor Manuel Albores Alcázar en julio de 1996 como chofer de transporte escolar, Karla Fabiola Albores Alcázar se integró en la docencia en 1998, Martha Patricia Albores Alcázar en el departamento de cobranza en 1999.

En el año 2002, Víctor Manuel Albores Alcázar formó el Grupo Educativo Albores Alcázar S.C. para darle un nuevo rumbo y sentido empresarial al negocio familiar y en el año 2004 funda la Universidad Del Sureste.

La formación de nuestra Universidad se da principalmente porque en Comitán y en toda la región no existía una verdadera oferta Educativa, por lo que se veía urgente la creación de una institución de Educación superior, pero que estuviera a la altura de las exigencias de los jóvenes que tenían intención de seguir estudiando o de los profesionistas para seguir preparándose a través de estudios de posgrado.

Nuestra Universidad inició sus actividades el 18 de agosto del 2004 en las instalaciones de la 4ª avenida oriente sur no. 24, con la licenciatura en Puericultura, contando con dos grupos de cuarenta alumnos cada uno. En el año 2005 nos trasladamos a nuestras propias instalaciones en la carretera Comitán – Tzimol km. 57 donde actualmente se encuentra el campus Comitán y el corporativo UDS, este último, es el encargado de estandarizar y controlar todos los procesos operativos y educativos de los diferentes campus, así como de crear los diferentes planes estratégicos de expansión de la marca.

Misión

Satisfacer la necesidad de Educación que promueva el espíritu emprendedor, aplicando altos estándares de calidad académica, que propicien el desarrollo de nuestros alumnos, Profesores, colaboradores y la sociedad, a través de la incorporación de tecnologías en el proceso de enseñanza-aprendizaje.

Visión

Ser la mejor oferta académica en cada región de influencia, y a través de nuestra plataforma virtual tener una cobertura global, con un crecimiento sostenible y las ofertas académicas innovadoras con pertinencia para la sociedad.

Valores

- Disciplina
- Honestidad
- Equidad
- Libertad

Escudo



El escudo del Grupo Educativo Albores Alcázar S.C. está constituido por tres líneas curvas que nacen de izquierda a derecha formando los escalones al éxito. En la parte superior está situado un cuadro motivo de la abstracción de la forma de un libro abierto.

Eslogan

“Mi Universidad”

ALBORES



Es nuestra mascota, un Jaguar. Su piel es negra y se distingue por ser líder, trabaja en equipo y obtiene lo que desea. El ímpetu, extremo valor y fortaleza son los rasgos que distinguen.

Doctorado En Educación

Objetivo de la materia:

Conocer la aproximación conceptual a la idea de evaluación o proceso de evaluación, cuáles son sus componentes y principales elementos a tener en consideración en la construcción de una idea más comprensiva de la evaluación.

Criterios de evaluación:

No	Concepto	Porcentaje
1	Trabajos en plataforma Educativa	60%
2	Examen	40%
Total de Criterios de evaluación		100%

INDICE

Unidad I

EVALUACIÓN DEL APRENDIZAJE EDUCATIVA

1.1.- Ámbitos de evaluación educativa.

1.2.- Modelos de evaluación.

1.3.- Planeación del proceso de evaluación de instituciones educativa.

1.4.- Instrumentos para la evaluación de una institución educativa.

Actividad para semana 1:

Se realizará un cuadro sinóptico: los conceptos principales se presentan de manera adecuada, clara, pertinente y precisa, lo cual demuestra lectura y análisis del tema. Incluye todos los conceptos secundarios importantes que representa la información principal del tema-

I.1 LA EVALUACIÓN EDUCATIVA

EDUARDO L. DE LA GARZA VIZCAYA

La naturaleza de la evaluación educativa

La Revista Mexicana de Investigación Educativa dedica la sección temática de este número a la evaluación educativa. La evaluación, como concepto integrante de las políticas públicas, ha estado presente en el contexto educativo mexicano, en los últimos veinte años. El término a estado vinculado con otro concepto: el de la calidad de la educación. A la evaluación se la concibe como una actividad indispensable y previa a toda acción conducente a elevar el nivel de la calidad de la educación. La evaluación así formulada constituye un momento de la planeación, entendida ésta como una acción racional dotada de propósito. Se identifica a la evaluación con la última etapa del proceso natural del conocimiento que concluye con la emisión de juicios informados, proceso que antecede a las decisiones y a la acción humana.

En su forma más simple, la evaluación conduce a un juicio sobre el valor de algo y se expresa mediante la opinión de que ese algo es significativo. Se llega a este juicio calificando qué tan bien un objeto reúne un conjunto de estándares o criterios. Así, la evaluación es esencialmente comparativa. Supone la adopción de un conjunto de estándares y la especificación del grupo contra el cual el objeto es comparado. El objeto puede ser calificado como “bueno” o “malo”, cuando la referencia es la totalidad de los objetos o el objeto promedio del

grupo; o bien como “mejor” o “peor”, si es comparado con un subconjunto particular del grupo. La evaluación, en esencia, supone adoptar un conjunto de estándares, definirlos, especificar el grupo de comparación y deducir el grado en el cual el objeto alcanza los estándares. Una vez realizado lo anterior, el evaluador está en posibilidad de hacer, en un segundo momento, un juicio sobre el valor del objeto evaluado.

Es preciso notar que no todas las evaluaciones son racionales y explícitas; los estudios de caso revelan los estándares, comparaciones y juicios de manera implícita e intuitiva, y algunos buscan entender más que juzgar. El juicio de valor puede ser expresado mediante palabras ordinarias de tipo descriptivo y no necesariamente mediante calificativos de “bueno” o “malo”. Los ejemplos y anécdotas en una descripción pueden servir para este propósito, ya que el juicio de valor reside en la forma en que las palabras son usadas.

El proceso de evaluación real es complejo y no se ajusta a una mera deducción a partir de ciertos estándares; más bien se trata de un proceso dialéctico entre principios abstractos y ejemplos concretos, en donde quien evalúa verifica sus juicios contra aquellas situaciones particulares en las que se siente confiado de conocer sus resultados. Aún más, los estándares pueden estar en conflicto unos con otros, y será necesario balancearlos así como establecer precedencias relativas entre ellos, algunas veces de manera intuitiva. Cuando el evaluador es distinto al que toma las decisiones, hay otro factor que

interviene: la credibilidad. La preocupación del tomador de decisiones no es sólo atender al grupo de comparación y a los estándares, y cómo éstos son pesados relativamente entre sí, sino también si los hechos son reales y si están presentados por el evaluador en forma correcta.

Normalmente, la evaluación educativa supone la existencia de acuerdos comunes: juicios previos consensuados, sobre lo que es bueno o conveniente para todos como propósito de las acciones educativas. Este es el espacio que permite el disenso y la controversia sobre el sentido, el propósito y el uso de la evaluación en la educación. Es, asimismo, el que permite adjudicarle un carácter ideológico. La evaluación no es meramente un proceso técnico, neutral e inocuo; por el contrario, como toda acción social posee una dimensión política y, en su carácter instrumental, obedece a una lógica de control. Estos aspectos constitutivos de la evaluación están presentes, lo mismo en una actividad tan antigua y concomitante a los procesos de educación como la evaluación en el aula, como en las recientes políticas gubernamentales identificadas con procesos de meta gestión, los cuales condicionan y vinculan recursos económicos a los resultados de la evaluación.

La evaluación asumida como un campo particular de la investigación educativa es reciente. Se ha cuestionado incluso si los proyectos de evaluación pueden considerarse investigación educativa. Se arguye que el mismo carácter instrumental de la evaluación contraviene el sentido desinteresado de la

obtención de nuevo conocimiento que subyace a la lógica de la ciencia. Lo cierto es que, en un sentido amplio, el propósito de los proyectos de evaluación es la obtención de conocimiento, si bien se busca que éste sea útil para la toma de decisiones. La mayor parte de los estudios de evaluación requiere, asimismo, diseños metodológicos complejos y debe satisfacer normas que justifican la validez y confiabilidad del conocimiento que obtiene. En ello, las evaluaciones no se apartan de la parsimonia inherente a las investigaciones científicas.

Por otra parte, el fenómeno educativo, como el hombre mismo, es complejo y su conocimiento es siempre parcial e idiosincrático, pues está constreñido a la condición temporal de sus actores y posee una dimensión inevitablemente subjetiva. En este sentido, cualquier pretensión “científica” de las investigaciones educativas debe ser matizada, dado el carácter peculiar de su objeto de estudio. En lo que respecta a los estudios de evaluación, en los últimos años, se ha generalizado su aceptación como una rama de la investigación aplicada en educación.

1.1.1 El origen y evolución de la evaluación educativa

La evaluación educativa tiene su origen en los primeros años del siglo XX, en Estados Unidos. Está influida por las ideas de progreso, de la administración científica y por la ideología de la eficiencia social. El modelo de los objetivos de

aprendizaje impactó en los años treinta el desarrollo curricular y su evaluación. Para los sesenta era una tecnología poderosa que, junto con los diseños experimentales, constituían la metodología tradicional de la evaluación educativa. Evaluación, entonces, era sinónimo de medición, prueba o examen. Con la expansión educativa en las décadas de los cincuenta y sesenta surgieron las preocupaciones por la rendición de cuentas y la relación costo-beneficio, lo que condujo al empleo de sistemas de información y de procedimientos para monitorear los programas como parte de la administración social. A finales de los sesenta, se hace la crítica del modelo tradicional de evaluación y, en la siguiente década, surgen nuevas formas para evaluar las innovaciones educativas. Las críticas expresadas por evaluadores connotados como Atkin (1963), Cronbach (1963, 1975), Scriven (1967), Stake (1967), Parlett y Hamilton (1972) o Stenhouse (1975), señalaban que el modelo de objetivos limitaba la reforma curricular y trivializaba la naturaleza del aprendizaje. Algunos de ellos abogaban por una evaluación que permitiera entender el proceso de la innovación y no sólo emitir juicios acerca de su efectividad. Se decía que las evaluaciones no respondían a las preguntas de los educadores. Estos evaluadores enfatizaban la utilidad de la evaluación en el mejoramiento de los cursos y no en la obtención de juicios terminales. Destacaban el carácter multidimensional de cualquier innovación educativa, el cual no podía alcanzarse con los exámenes. Se decía que el modelo impedía juzgar el valor de

las metas propuestas, pues separaba los hechos fácticos de los valores implícitos. Alguno señalaba que conducía a una visión equivocada de la educación, ya que su pretensión de determinar las conductas de los estudiantes contravenía el propósito de la verdadera educación: el desarrollo de la autonomía personal, cuyos resultados son, en sí mismos, impredecibles. El diseño experimental, por otra parte, había mostrado que con frecuencia se daba un conflicto entre la confiabilidad y la relevancia de los resultados.

En la evolución subsiguiente de la evaluación educativa fueron relevantes cuatro conferencias realizadas en Cambridge, Inglaterra, a partir del inicio de los setenta, con el objeto de explorar acercamientos no tradicionales a la evaluación educativa y establecer algunas guías para su desarrollo futuro. Se buscaba justificar prácticas metodológicas y funciones políticas alternativas para la evaluación. El cambio de paradigma abrevaba en la tradición de las investigaciones humanistas, fuera de la psicometría, el experimentalismo y las encuestas sociales.

En la primera de ellas, llevada a cabo en 1972 (Hamilton, 1977), se estableció que la evaluación debería: a) responder a las necesidades y perspectivas de diferentes audiencias, b) iluminar los procesos complejos organizacionales, de enseñanza y de aprendizaje, c) ser relevante para las decisiones públicas y profesionales por venir, y d) reportarse en un lenguaje accesible a las audiencias. Recomendaba también que fueran usados datos proporcionados por

la observación, cuidadosamente validados (algunas veces en sustitución de datos proveídos por preguntas y exámenes); que la evaluación se diseñara de tal forma que fuera lo suficientemente flexible para responder a eventos no anticipados (que su foco fuera obtenido de manera progresiva, en lugar de con un diseño preordenado), y que la posiciones de valor del evaluador, tanto si fueran resaltadas como constreñidas por el diseño, se hicieran evidentes a los patrocinadores y a las audiencias de las evaluaciones.

Las tres restantes conferencias se llevaron a cabo en 1975 (Simons, 1980), cuyo foco fueron los métodos de la investigación mediante estudios de caso; en 1979, para examinar el uso de los métodos naturalistas, y en 1987 para considerar la evaluación del desempeño y diferentes formas de evaluar a maestros y alumnos.

Para 1979, había adquirido importancia política la rendición de cuentas y se hacía un énfasis excesivo en la administración. Para entonces, la evaluación educativa compartía los problemas de la investigación vinculada con las políticas. Entre los problemas detectados en la tercera conferencia se encontraban: la inadecuación de las teorías sobre políticas y los procesos para establecerlas; el énfasis usual en la evaluación de programas en cuanto opuesto a la evaluación de las políticas; el deterioro en la relación entre evaluadores y sus patrocinadores, clientes y otras audiencias, con frecuencia debido a la

abierta naturaleza política de la innovación; la dificultad de hacer coincidir las necesidades de información de los tomadores de decisiones y el servicio a audiencias múltiples dentro de la escala de tiempo establecida por los administradores de los programas; una historia de productos rechazados, especialmente por los administradores quienes esperaban reportes que simplificaran sus tareas en lugar de hacerlas más complejas; la dificultad de mantener la independencia; la vulnerabilidad personal e institucional vinculada con las formas de evaluación naturalista las cuales emplean retratos de las personas y descripciones amplias de los climas y contextos institucionales, y la naturaleza conservadora de esta misma investigación, la cual favorecía la descripción sobre el análisis.

A fines de la década setenta era claro que la evaluación había tenido poco impacto en las políticas educativas y que, en donde lo había tenido, había hecho más complejo el proceso de toma de decisiones. No se alcanzaba el propósito de proporcionar información útil y confiable.

1.1.2 El nuevo paradigma de la evaluación educativa

La nueva evaluación que surge de los cuestionamientos anteriores posee las siguientes características: a) es pragmática desde el punto de vista metodológico, b) responde a situaciones particulares: pone su atención en las preguntas, preocupaciones, problemas y necesidades de información de los

implicados y de los tomadores de decisión, c) hace a un lado el paradigma hipotético deductivo y acoge otro de decisiones que enfatizan métodos múltiples, acercamientos alternativos y el hacer coincidir los métodos de la evaluación con las preguntas y las situaciones específicas. Se llega a decir que el propósito más importante de la evaluación es responder a los requerimientos de información de las audiencias, particularmente, en formas que tomen en cuenta las variadas perspectivas de sus miembros. La intención es incrementar la utilidad de los resultados, aun sacrificando la precisión proporcionada por las mediciones.

Investigación naturalista ha sido el término genérico que fue usado para describir varios de los acercamientos alternativos a la evaluación que obtuvieron prominencia en los setenta como reacción a las formas tradicionales. Se han caracterizado como acercamientos que atienden la pluralidad de las audiencias y asumen el carácter político de la evaluación. Sus rasgos se expresan de la siguiente manera:

Comparados con los métodos clásicos, tienden a ser más extensivos (no necesariamente centrados en datos numéricos), más naturalistas (basados en la actividad del programa o proceso, más que en su intento o logro), y más adaptables (no constreñidos por diseños experimentales o preordenados). Asimismo, muy probablemente son sensibles a los diferentes valores de los participantes en el programa, suscriben métodos empíricos que incorporan

trabajo de campo etnográfico, desarrollan materiales de retroalimentación formulados en el lenguaje natural de los destinatarios, y desplazan el lugar de los juicios formales, del evaluador a los participantes (Hamilton, 1977).

Significan un compromiso de estudiar los programas en su contexto social, el uso de métodos cualitativos de indagación, tales como las entrevistas no estructuradas, la observación directa y la reconstrucción histórico dramática, y formas de reportar que permiten a los lectores generalizar por sí mismos, utilizando una generalización naturalista (Simona, 1987).

Dirigidos a audiencias no técnicas resaltando el uso del lenguaje ordinario, el razonamiento informal y cotidiano y el uso de argumentos para tratar de establecer la estructura de la realidad. Estos argumentos son: los ejemplos, las metáforas, las ilustraciones y las analogías (House, 1980).

1.1.3 La evaluación educativa en México en los últimos años

Las acciones y los estudios de evaluación han estado creciendo y se han consolidado en el medio mexicano en los últimos quince años. Las pruebas objetivas en gran escala, aplicadas a poblaciones del nivel básico, estudiantes y en menor medida maestros, recibieron un impulso importante a mediados de los noventa desde la Secretaría de Educación Pública y, recientemente, con la creación en 2002 del Instituto Nacional para la Evaluación de la Educación (INEE). La creación en 1994 del Centro Nacional de Evaluación

para la Educación Superior (CENEVAL) generalizó también el empleo de estas pruebas en el ámbito de la educación superior. En este nivel, ha cobrado importancia la evaluación de programas a través de pares a partir de la creación de los Comités Interinstitucionales para la Evaluación de la Educación Superior en 1991 y, desde 1994, de los organismos integrados en el Consejo para la Acreditación de la Educación Superior, en 2000. Últimamente, también en el ámbito del nivel superior, ha destacado el Programa Integral de Fortalecimiento Institucional (PIFI) iniciado en 2001, para impulsar la evaluación y planeación de las instituciones.

En el nivel de la educación básica es pertinente mencionar dos líneas más cultivadas actualmente por el INEE. La primera, es la evaluación cualitativa de las escuelas, la cual había sido iniciada previamente por la SEP en programas como el de Escuelas de Calidad. La segunda línea se refiere a la definición de indicadores educativos, aspecto también atendido previamente en la SEP. Desde su creación, el INEE ha fomentado la elaboración de investigaciones y estudios vinculados con sus tareas de evaluación; recientemente lo ha hecho mediante una convocatoria abierta a la comunidad de investigadores. También el CENEVAL ha sido un factor importante para impulsar la elaboración de estudios, pues ha auspiciado, desde 1995, la celebración de foros nacionales bianuales dedicados a la evaluación de la educación. El recuento anterior no pretende ser exhaustivo, sino sólo mostrar el creciente impulso de las

actividades de evaluación en nuestro medio, muchas de ellas derivadas de políticas gubernamentales que hacen eco a tendencias de orden internacional. Tampoco es aquí el lugar para hacer una relación de las investigaciones o estudios de corte evaluativo realizados en el país, lo cual rebasaría el propósito de esta introducción. Baste decir que constituyen un campo de investigación creciente y en proceso de maduración.

1.1.4 Las investigaciones en este número

A pesar de la afirmación anterior, la respuesta de la comunidad a la convocatoria para este tema fue reducida. Se presentaron sólo cinco trabajos, todos del medio mexicano. Cuatro de ellos fueron aprobados para integrar la sección que ahora introducimos. En él, el lector encontrará una muestra de las investigaciones educativas de corte evaluativo que se realizan actualmente en el medio mexicano. Tres de ellas son de naturaleza empírica. En éstas, se ha pedido expresamente a los autores reflexionar sobre la metodología utilizada. Consideramos, como debe haber quedado claro con las ideas anteriores, que las investigaciones de corte evaluativo se benefician al explicitar y justificar su método, en relación con su propósito.

El primer trabajo, “¿Aprobar o reprobar? El sentido de la evaluación en educación básica”, es una investigación documental que cuestiona, con base en experiencias y en la evidencia de estudios empíricos realiza-

contextos distintos al mexicano, la práctica de la reprobación en ese nivel educativo. En la medida en que se considera al éxito un indicador de la calidad educativa, esta discusión se hace pertinente. Los lectores familiarizados con la educación en los niveles superiores encontrarán también resonancia en algunas conclusiones que vinculan el abandono con la reprobación.

El segundo texto, “Evaluación del desempeño en lectura y escritura. Aportes empíricos a la noción de componentes lingüísticos en el cuarto grado de primaria”, se detiene en el debate sobre la evaluación del aprendizaje, su sentido y el uso de sus resultados. La investigación postula la necesidad de explicitar “una teoría que sustente tanto a los medios de medición utilizados como a las características y propiedades a evaluar” y aporta hallazgos significativos sobre la pertinencia de organizar en bloques el currículum de español en primaria.

El tercer artículo, “Evaluación de la docencia universitaria: una propuesta alternativa que considera la participación de los profesores”, es una investigación que privilegia el carácter formativo de la evaluación, es decir su utilidad para mejorar la calidad del mismo proceso educativo. El trabajo pretende contribuir a la construcción de un modelo de evaluación desde la práctica, “a partir del análisis de la vida cotidiana en las aulas”. El evaluador en este modelo, “acompaña” al docente como consultor y orientador, ayudándole en la reflexión sobre su propia práctica.

El cuarto trabajo, “Evaluación de la docencia y el contexto disciplinario: la opinión de los profesores en el caso de ingeniería y tecnología”, es un estudio que explora las dimensiones de la docencia que, a sugerencia de los propios docentes, deben incorporarse a un cuestionario de evaluación. La investigación identifica componentes relevantes de la enseñanza en la disciplina de estudio. Como en los dos trabajos anteriores, los autores reflexionan sobre el estado del arte de la investigación y las metodologías empleadas, en este caso para la evaluación de la docencia, y hacen aportaciones importantes sobre el uso de los cuestionarios de opinión de alumnos como recurso generalizado.

La entrega del número temático se enriquece con una colaboración en la sección de Aporte de discusión de dos investigadores, quienes –desde la experiencia de la evaluación en el contexto español– abogan por la necesidad de iniciar procesos de “meta evaluación”. Presentan una serie de criterios o parámetros que deben ser satisfechos por las evaluaciones. En su propuesta, el propósito de la meta evaluación es “avivar la reflexión para generar comprensión rica y relevante sobre los programas evaluados” y, al mismo tiempo, garantizar la calidad de la evaluación. Los autores destacan seis contextos “básicos” de toda evaluación que permiten explorar los elementos a ser considerados en la meta evaluación. Más adelante, se detienen en los aspectos de la validez y confiabilidad de las evaluaciones, y discuten su sentido

en los distintos modelos utilizados. Se puede estar de acuerdo o no con las posiciones de los autores y el tratamiento que hacen de los tópicos, pero resulta interesante el propósito, aunque no del todo exitoso, de sistematizar en una perspectiva única paradigmas cuantitativos y cualitativos que, hasta hace poco, se ubicaban en parcelas irreconciliables.

La colaboración termina discutiendo tópicos relevantes en los estudios de evaluación: su propósito, la generalización del conocimiento obtenido, la ética y el aprendizaje.

1.1.5 Los ámbitos de la evaluación educativa y algunas falacias frecuentes

La asociación entre la evaluación y la práctica educativa escolar parece muy directa. La sola mención de la palabra evaluación nos remite a escenas y situaciones reiteradas en nuestra propia experiencia como alumno en cualquier tramo de nuestra trayectoria educativa. Sin embargo, se puede reconocer que estas asociaciones siempre responden a un mismo patrón de evaluación, a estrategias relativamente similares entre sí, a la idea de alguna forma de control, a la idea de finalización de una etapa o ciclo y a un conjunto de consecuencias directa e indirectas, no siempre claramente establecida, derivadas de los resultados de la evaluación. Pero no siempre cabe considerar a la evaluación del mismo modo y adoptando formas tan poco variadas entre sí.

Con el propósito de avanzar en la construcción de un concepto más amplio sobre la evaluación educativa, en esta oportunidad parece adecuado detenerse en la consideración de un componente fundamental de cualquier proceso de evaluación: la correcta y precisa definición del ámbito de la evaluación, es decir el espacio o los límites dentro de los cuales tendrá lugar el proceso de evaluación. Esta es una definición complementaria e imprescindible a la pregunta sobre el objeto de la evaluación y sobre el propósito de la misma.

La pregunta por el ámbito de la evaluación determina la necesidad de precisar el alcance de la misma y el conjunto de herramientas e instrumentos que por su adecuación se seleccionarán para llevar a cabo dicho proceso, es decir, se definen el alcance y el marco metodológico a utilizar.

En el marco de la problemática educativa es posible sostener al menos la diferenciación clara entre cuatro ámbitos posibles para el desarrollo de una propuesta de evaluación: de los aprendizajes; de las instituciones; de los proyectos y programas; del sistema educativo.

La evaluación de los aprendizajes reconoce como ámbito o espacio típico aquello que sucede en el aula, demanda estrategias metodológicas precisas

destinadas a registrar indicios claros sobre un proceso tan complejo e individual como es el aprendizaje. En este terreno si bien se encuentra una práctica frecuente no se ha desarrollado gran variedad de estrategias,

Instrumentos y herramientas. Se observa una tendencia sostenida a la reiteración en el uso de los mismos y la exigencia de adecuación de las mismas herramientas para diferentes propósitos evaluativos, con diferentes consecuencias para los alumnos y con el objetivo de consignar el rendimiento o el alcance de diferentes tipos de aprendizajes.

La evaluación de las instituciones reconoce un ámbito muy claro y muy complejo como es el de la institución escolar, en este caso demanda estrategias metodológicas muchas veces próximos al marco general de lo que se conoce como análisis institucional aplicado a otros espacios organizacionales, pero con el requerimiento básico de reconocer la especificidad de la escuela como forma particular de institución. La evaluación institucional requiere de una definición clara de los propósitos y los alcances de la misma al igual que de las limitaciones que se reconocen desde un inicio. Otro elemento fundamental en la definición de una propuesta de estas características es quién o quiénes son los responsables de la evaluación y quienes intervienen de uno u otro modo. En este sentido se abre un abanico interesante de combinaciones entre formas de

evaluación externa e iniciativas de autoevaluación que una vez más obedecen a propósitos diferentes e implican desarrollos metodológicos también diversos.

Evaluación de programas y proyectos, en este caso el ámbito es el delimitado por un conjunto organizado y coherente de decisiones que toman cuerpo en un proyecto educativo, en un programa de trabajo. Este terreno de los programas o proyectos nuevamente es muy amplio y es evidente que cualquier propuesta de evaluación en este sentido tiende a incluir componente ligados a las estrategias conocidas de evaluación de proyectos y en especial de proyectos sociales. Sin embargo, nuevamente el terreno educativo requiere de algunas especificidades y no siempre los criterios adoptados para cualquier tipo de proyecto social resultan satisfactorios.

Es necesario reconocer en este terreno la diferencia que implica la evaluación de proyectos o programas educativos originados en diversas organizaciones o entidades de aquellos proyectos o programas que forman parte de una política educativa, es decir cuyo promotor es el Estado. En este sentido esta evaluación de programas y proyectos se acerca considerablemente a una evaluación de la política educativa o de las políticas educativas. Volviendo la mirada hacia los requerimientos metodológicos es necesario poder precisar cuál o cuáles serán

los ejes centrales de la evaluación del proyecto, programa o política: el diseño, la gestión y ejecución, los resultados, el impacto, etc.

Evaluación del sistema educativo, en este caso el ámbito es el más amplio posible y se refiere a la búsqueda de respuestas sobre el funcionamiento más o menos adecuado de la totalidad o un segmento del sistema educativo en su conjunto con todas las complejidades que esto supone. La consideración de la evaluación en este ámbito supone la inclusión de algunos componentes vinculados típicamente con otros ámbitos, pero donde la lectura, el análisis y

el uso de la información suponen juicios de valor sobre las características distintivas del sistema educativo en su conjunto. En este caso los propósitos se asocian con la toma de decisiones a nivel macro y la definición de políticas educativas y líneas de acción a nivel nacional o subnacional y no se refieren a situaciones particulares o individuales en ningún caso. Los requerimientos metodológicos son bien complejos e implican las decisiones respecto al alcance y los propósitos de la evaluación. Cualquier evaluación en este ámbito supone complejos dispositivos técnicos y logísticos a la vez que fuertes consecuencias e impacto en el conjunto de los actores involucrados en el sistema educativo y la opinión pública en general.

Esta distinción de los ámbitos de la evaluación educativa puede parecer muy obvia pero no lo es en la práctica actual de la evaluación educativa y menos aún cuando llega el momento de extraer conclusiones una vez desarrollado el proceso de evaluación.

Con frecuencia se producen algunas falacias que no contribuyen en nada en la construcción de una idea más amplia y útil de la evaluación y que en buena parte son derivadas de la escasa o nula identificación precisa del ámbito en que se desarrolla la evaluación, algunas de ellas son las siguientes.

- Desarrollar un proyecto de evaluación propio de un ámbito y extender el análisis y las conclusiones a otro ámbito diferente. Por ejemplo, es frecuente sostener acciones de evaluación centradas en la evaluación de los aprendizajes y extender dichas conclusiones como suficientes para explicar el funcionamiento institucional o del sistema educativo.
- Utilizar el marco metodológico, las herramientas y los instrumentos propios de un ámbito específico para desarrollar acciones de evaluación dentro de otro ámbito. Tienen lugar frecuentemente una suerte de abuso de las herramientas e instrumentos, un ejemplo simple lo ofrecen las diferentes formas de test o exámenes regularmente escritos y de estructura cerrada – de elección múltiple por ejemplo – que están destinados a la evaluación de

determinados tipo de aprendizajes, en aplicaciones masivas y con propósitos bien explícitos, sin embargo con mucha frecuencia se intenta utilizar estos instrumentos o criticar su utilización como los únicos posibles para la evaluación de los aprendizajes, dejando de lado la consideración del ámbito específico en que esta tiene lugar.

- Diseñar proyectos de evaluación que suponen la consideración de más de un ámbito simultáneamente sin reconocer las diferencias existentes y necesarias entre los mismos. Esto conduce a cierta pobreza de los resultados del proceso de evaluación ya que en su desarrollo habitualmente se producen estas definiciones – ámbito, objeto, propósito – explícita o implícitamente sin embargo la escasa precisión del ámbito de desarrollo de la acción de evaluación en forma previa trae como consecuencia un planteo confuso, poco constructivo y el escaso aprovechamiento de las instancias de evaluación para la generación de aprendizajes relevantes alrededor del objeto de evaluación.

En síntesis, la preocupación por los ámbitos de la evaluación educativa no es más que aquella por incrementar la precisión, y por ende la relevancia, de los procesos de evaluación. Se aspira a que las acciones de evaluación educativa resulten efectivas en términos de mejor, es decir, produzcan información significativa sobre aquellos procesos que tienen lugar en el espacio educativo y que por su complejidad no resultan sencillos de abordar. Se intenta así, advertir

sobre la necesidad de diseñar proyectos de evaluación que atiendan a las especificidades del objeto de evaluación y que desarrollen simultáneamente herramientas metodológicas variadas y adecuadas a la vez que capacidades y competencias específicas en quienes llevan a cabo esta labor. Solo de este modo la evaluación como tal significa un aporte relevante en términos de acciones de mejora.

1.2 MODELOS DE EVALUACIÓN.

ENFOQUES Y MODELOS DE EVALUACIÓN DEL E-LEARNING

El objetivo es presentar el estado de la cuestión sobre la evaluación del e-learning. Con este propósito se muestran los diferentes enfoques evaluativos, así como diferentes modelos, herramientas y experiencias encaminadas a determinar la calidad de la formación virtual, o e-learning. Se destacan dos enfoques principales, la evaluación de enfoque parcial, que enfatiza aspectos diversos del e-learning (los materiales, los recursos tecnológicos, la docencia, etc.) y la evaluación de enfoque global, que utiliza modelos de la gestión de la calidad y la práctica del benchmarking.

E-learning es un producto más de los generados por la sociedad de la información y la era digital, que cobra una especial importancia en el marco de los nuevos modelos de enseñanza/aprendizaje y del aprendizaje a lo largo de toda la vida en convergencia con las posibilidades que las tecnologías de la información y la comunicación ofrecen a las aplicaciones educativas.

Desde el punto de vista conceptual e- learning es un término susceptible de diferentes definiciones y a menudo intercambiable por otros: formación on-line, cursos on-line, formación virtual, tele formación, formación a distancia, campus virtual... En sentido literal, del inglés, significa aprendizaje electrónico, el aprendizaje producido a través de un medio tecnológico-digital.

Rosenberg (2001) lo define como el uso de las tecnologías basadas en internet para proporcionar un amplio despliegue de soluciones a fin de mejorar la adquisición de conocimientos y habilidades. El mismo autor establece tres criterios que se han de cumplir para poder aplicar correctamente el término (Rosenberg, 2001: 28-29):

- a) que se produzca en red, lo que permite una actualización inmediata, almacenamiento y recuperación, distribución y capacidad de compartir los contenidos y la información,
- b) que llegue al usuario final a través de un ordenador, utilizando estándares tecnológicos de Internet,
- c) que esté centrado en la visión más amplia de soluciones para el aprendizaje que van más allá de los paradigmas tradicionales de la formación.

Sin embargo estos criterios son fuertemente restrictivos. El “aprendizaje electrónico” no sólo se produce a través de estándares tecnológicos de Internet, las plataformas ad hoc son un elemento cada vez más importante; y los materiales off line o descargables de apoyo al resto de la formación son un

componente primordial mientras no mejoren las infraestructuras de las comunicaciones en la propia red.

Desde su irrupción en el mundo educativo y formativo, el e-learning ha generado importantes expectativas no sólo de carácter pedagógico, sino también de carácter social y económico, lo que unido al creciente interés por la calidad educativa (González, 2000:53) en cualquiera de sus manifestaciones y ámbitos, hace que se imponga la necesidad de desarrollar modelos de evaluación adecuados al objeto y a los distintos contextos en los que se produce.

I. Tendencias en la evaluación del e- learning

La inquietud por evaluar el e-learning está dando lugar a importantes iniciativas y experiencias a nivel mundial encaminadas a establecer estándares que permitan certificar su calidad.

Aunque los mecanismos de parametrización de la calidad varían en función tanto del contexto (Tait, 1997), como del propio concepto de calidad (Harvey y Green, 1993), hasta el momento se puede hablar de dos grandes tendencias en relación a las prácticas para evaluar la calidad de las instituciones y de los proyectos que utilizan el e-learning como actividad formativa con entidad propia (Sangrà, 2001). El objetivo se centra sobre todo en buscar criterios e indicadores específicos que den respuesta a las preguntas que se plantea la evaluación de la calidad de la formación en entornos específicos, con medios

específicos y dirigida a personas con un perfil diferente al del alumno tradicional (en el caso de las universidades).

Enfoque parcial. Centrado principalmente en alguno de los siguientes aspectos:

La actividad formativa

Los materiales de formación

Las plataformas tecnológicas

La relación coste/beneficio

Enfoque global. Se distinguen dos tendencias:

Los sistemas de evaluación centrados en modelos y/o normas de calidad estándar y calidad total

Sistemas basados en la práctica del benchmarking

I.I. La evaluación de enfoque parcial

Es la evaluación centrada en alguno de los elementos considerados de mayor interés dentro de una solución e-learning.

I.I.I. Evaluación de la actividad formativa

Es el proceso orientado a evaluar una acción concreta de formación, como puede ser un curso online, de mayor o menor duración. La finalidad de esta evaluación se orienta hacia tres aspectos básicamente: comprobar el nivel de cumplimiento de los objetivos educativos, mejorar la propia acción formativa y determinar el retorno de la inversión realizada.

Belanger y Jordan (2000:187) identifican tres modelos principales en la evaluación de acciones formativas (adaptados algunos de la formación tradicional). Modelos que ponen el énfasis bien en la evaluación diagnóstica, antes de introducir la acción formativa, bien en la evaluación final, una vez se ha producido la formación.

1.2.1 Modelo Sistémico de Vann Slyke et al. (1998)

El modelo provee de un conjunto de variables que interactúan como factores predictores del éxito de la acción formativa on-line. Estas se concentran en las siguientes:

Características institucionales

Las características institucionales están relacionadas con la capacidad de la organización para implementar acciones de e-learning, tales como los objetivos de la institución, la infraestructura de soporte a la acción, la capacidad económica.

Características de los destinatarios de la formación

Estas características están relacionadas con los intereses, expectativas y habilidades de los estudiantes (autosuficiencia, gestión personal del tiempo, dominio del ordenador y actitud hacia la tecnología, capacidad para la resolución de problemas). Es el único modelo que presenta la variable “características del alumnado” como factor de éxito o fracaso de la formación on-line, aunque son varios los autores que enfatizan las diferencias individuales

de los usuarios como elemento importante predictor del éxito de la formación virtual (Richardson, 2001; Oliver, 1998; Ramussen y Davidson, 1996).

Características del curso

Las características del curso tienen que ver con la capacidad del sistema de e-learning en relación a las necesidades y metodologías de enseñanza-aprendizaje para el curso. Por ejemplo, si el curso requiere una metodología basada en el trabajo colaborativo el entorno virtual debe poder facilitararlo.

Características de la formación a distancia

Estas características se refieren al hecho de la necesidad de crear nuevos modelos de acomodación de los usuarios a los nuevos entornos, de forma que se asegure su tranquilidad, confort y facilidad de aprendizaje.

Estas variables deben ser estudiadas para la implementación de acciones formativas on-line de forma que se adapten al alumnado, al profesorado, a la institución y a la sociedad. Así por ejemplo los estudiantes pueden acabar recibiendo una alta interacción con el profesor si lo necesitan, la institución puede incrementar la productividad entre los docentes o a nivel social puede mejorarse el acceso a la educación, la calidad de vida, la fuerza del trabajo, etc.

Modelo de los cinco niveles de evaluación de Marshall and Shriver (en McArdle, 1999)

Este modelo se centra en cinco niveles de acción orientados a asegurar el conocimiento y competencias en el estudiante virtual.

Docencia

Este nivel se centra en la capacidad del docente en la formación on-line para proyectarse a través del medio tecnológico (el correo electrónico, el chat, el aula virtual, ...), haciendo uso de habilidades comunicativas adecuadas a ese entorno, tales como la claridad en la redacción de los mensajes, intervención frecuente en el aula virtual, inmediatez y eficacia en las respuestas a los mensajes del alumnado, apropiación adecuada de los recursos que provee el entorno tecnológico

Materiales del curso

La evaluación de los materiales debe ser realizada por el alumnado con relación al nivel de dificultad, pertinencia, interés o efectividad.

Curriculum

Los contenidos o el currículum del curso deben ser evaluados con un nivel elevado de análisis y por comparación con otros currícula.

Módulos de los cursos

La modulación es una característica de los cursos on-line que debe igualmente ser valorada en relación a su estructura y orden.

Transferencia del aprendizaje

Este último nivel persigue determinar el grado en el que el curso on-line le permite a los participantes transferir los conocimientos adquiridos al puesto de trabajo.

El modelo combina diferentes elementos del acto educativo, pero pone un especial énfasis en el docente, como agente dinamizador de la formación en entornos virtuales. De hecho asistimos a un renovado interés por la calidad docente (Mateo, 2000a), que se está convirtiendo en un factor estratégico de primera magnitud. En el entorno virtual el docente se ha de situar en un nuevo espacio formativo como guía y acompañante del protagonista del aprendizaje, el participante, siendo la interacción la base para el desarrollo formativo (Duart, 2001a). Aunque los estándares sobre los niveles de actuación del profesorado son distintos en la formación on-line, muchos de los parámetros considerados en los modelos de evaluación del docente presencial (Mateo, 2000b) pueden ser los mismos.

Modelo de los cuatro niveles de Kirkpatrick (1994)

El modelo de Kirkpatrick ha sido y es ampliamente utilizado en la evaluación de acciones formativas tradicionales, y en la actualidad son varios los autores que recomiendan su adaptación y uso en el e-learning (Rosenberg, 2001, Mantyla, 2000, Belanger y Jordan, 2000).

El modelo está orientado a evaluar el impacto de una determinada acción formativa a través de cuatro niveles: la reacción de los participantes, el aprendizaje conseguido, el nivel de transferencia alcanzado y finalmente el impacto resultante.

Reacción

Es sin duda el tipo de evaluación más utilizada en la mayoría de cursos de formación. Puede realizarse a través de un cuestionario de opinión, o de forma más cualitativa mediante grupos de discusión. En el e-learning cobra un interés especial si se tiene en cuenta que es el único feed-back sobre la reacción de los usuarios con el que cuentan los gestores del curso (frente a distintos elementos, como el/la docente, los materiales, los contenidos, el entorno, el aprendizaje, la transferencia o la percepción del impacto de la formación recibida).

Aprendizaje

Esta evaluación (en su versión formativa y/o sumativa) persigue comprobar el nivel de conocimientos y habilidades adquiridos por el alumnado a través de test o pruebas de rendimiento validadas y fiabilizadas (Mantyla, 2000: 267). Kirkpatrick (1999) recomienda el uso de una metodología cuasiexperimental como estrategia para poder establecer de forma más objetiva la efectividad del curso. Otros autores en cambio consideran que más que buscar la efectividad del curso esta evaluación debe utilizarse como método de retroalimentación para mejorarlo (Rosenberg, 2001). En cualquier caso el reto en la formación online con relación a la evaluación de los aprendizajes está en configurar estrategias y sistemas de validación que no requieran de la presencialidad.

Transferencia

La evaluación de la transferencia consiste en detectar si las competencias adquiridas con la formación se aplican en el entorno de trabajo y si se mantienen a lo largo del tiempo (mejor desempeño de la tarea, más rapidez, menos errores, cambio de actitud, etc.). Evaluar la transferencia permite demostrar la contribución de la formación a la mejora de las personas y los beneficios que aporta a la organización para determinar más tarde su impacto y rentabilidad. A pesar de ser una evaluación crucial en la formación empresarial, sólo un 62% de las organizaciones la utilizan, según investigaciones realizadas (EPISE, 2000, Pineda et al., 1999). Los instrumentos o estrategias más utilizados son la observación, las entrevistas a los supervisores y pares y la autoevaluación de los participantes (Pineda, 2002).

Impacto Igual que sucede con el nivel anterior también esta evaluación es utilizado sobre todo por las empresas. Aunque tradicionalmente la evaluación del impacto o de los resultados se ha basado en criterios económicos (demostrar un mayor número de ventas, mayor productividad, menos errores, calidad de servicio, menos reclamaciones, ...), estos no siempre son posibles de probar, por lo que Rosenberg (2001: 223) afirma que “Es mejor conformarse con la evidencia que reclamar pruebas. Si al preguntar al gerente de la empresa y a todos los supervisores si la introducción reciente de soluciones e-learning genera una mayor productividad, la gran mayoría de ellos contestan que sí, ello es una buena evidencia de que existe correlación entre las soluciones e-learning

y la medición de los negocios”. El objetivo de la evaluación también podría ser determinar hasta qué punto la falta de formación puede llegar a tener un impacto perjudicial en la organización. De igual forma hay que advertir que el impacto de la formación no sólo se produce a nivel económico, el conocimiento producido, la capacidad de innovación que genera o la fidelización de las personas empleadas es también importante.

El problema general que presentan todos los modelos referenciados en este epígrafe es que ninguno de ellos esclarece con nitidez los indicadores de evaluación, ni los estándares de valoración, ni los procesos y formas de obtención de evidencias de cada uno de los elementos evaluados.

1.2.2 Evaluación de los materiales

La calidad de los materiales formativos cobra una significación especial en la formación no presencial, al ser el instrumento principal de transmisión básica de conocimientos del que dispone el alumnado. De ahí que su evaluación se haya convertido en una de las evaluaciones a las que se le han dedicado mayores esfuerzos.

Los materiales utilizados en el e-learning pueden ser textuales, hipertextuales (o hipermedia) o multimedia, y estar diseñados para su uso tanto on-line como off-line

En el ámbito de la evaluación destaca la línea de evaluación de materiales multimedia, también llamada evaluación de software educativo, con una larga

tradición desde la aparición de las primeras aplicaciones informáticas educativas (EAO). Bajo parámetros de la evaluación de programas, las experiencias relacionadas con la evaluación de materiales multimedia han venido centrándose según el objetivo de la evaluación y el agente evaluador en la evaluación de necesidades, del input, del proceso, del producto y/o de los resultados (tabla 1). Pero de forma particular destaca el despliegue desarrollado para la evaluación del producto orientada a certificar la calidad de los materiales y/o facilitar la toma de decisiones en su selección.

Cabero (2001:451-455) identifica tres tipos de evaluación con respecto a los medios tecnológicos en general: la evaluación del medio en sí (características del medio), la evaluación comparativa con otro medio y la evaluación didáctico curricular (el comportamiento del medio en el contexto de enseñanza-aprendizaje); y tres agentes evaluadores: los productores, los expertos (en contenidos, diseño instructivo), y los usuarios.

Existen numerosas iniciativas encaminadas a diseñar instrumentos de medidas estándar de calidad para la evaluación de los materiales tecnológicos educativos. Entre ellas destacan:

Por su magnitud: que reúne a empresas americanas multinacionales informáticas e instituciones educativas para definir estándares tecnológicos; o the Promoting Multimedia Acces Education and Training in European Society (<http://www.perseus.tufts.edu>), un proyecto orientado a definir la calidad

formativa de los materiales multimedia. De forma más modesta, pero también destacable, encontramos, a modo de ejemplo: el Modelo Sistémico de Calidad de Software (Mendoza et al., 2001) orientado a determinar la calidad del software en baja, media y elevada; el proyecto E-CumLaude (Rodríguez et al., 2001) para certificar la calidad de los materiales multimedia educativos centrado en criterios de calidad de mínimos y de excelencia; el método de Evaluación de Materiales Educativos Computarizados basado en el juicio de expertos (Galvis, 2000); o la escala de Catalogación y Evaluación Multimedia SAMIAL (Navarro, 1999) centrada en distintas categorías y orientada a establecer la calidad del material en excelente, alta, correcta y baja.

<i>Funciones de la evaluación</i>	<i>Objetivos de la evaluación y agentes evaluadores</i>	<i>Criterios para la evaluación</i>	<i>Instrumentos de obtención de información</i>
<i>De necesidades</i>	Aportar información sobre las mejoras que puede introducir el nuevo material Productores	Estudio de otros materiales con objetivos similares	Análisis cualitativo
<i>Del input</i>	Determinar las capacidades para la realización del material Productores	Software y hardware (de desarrollo y del usuario) Programación	Análisis cualitativo
<i>Del proceso</i>	Corregir y perfeccionar el material durante su desarrollo Productores, usuarios potenciales	Control de los contenidos Control funcional	Protocolo de evaluación
<i>Del producto</i>	Juzgar la calidad del material en sí una vez acabado y/o tomar decisiones para su utilización Expertos externos, seleccionadores, productores	Ámbito disciplinario Ámbito didáctico Ámbito tecnológico	Protocolo de evaluación, entrevistas, grupos de discusión

<i>De los resultados</i>	Determinar los usos y el funcionamiento del material en un contexto de enseñanza-aprendizaje Comparar el material con otros en cuanto a efectividad Usuarios reales, investigadores, seleccionadores	Opinión de los usuarios Utilidad Relación coste-beneficios	Cuestionarios, escalas de actitud, pruebas de aptitud
--------------------------	--	--	---

Frente a instrumentos estructurados como los que acabamos de presentar aparecen propuestas más flexibles que conducen a la identificación de una serie de criterios para que sea el usuario, seleccionador o evaluador el que determine la ponderación de los mismos y la reflexión sobre la pertinencia de su presencia o ausencia. En esta línea se encuentra el instrumento promovido por la European Academic Software Award (Baumgartner y Payr, 1997) desarrollado en torno a doce criterios: exactitud, relevancia, cobertura, interacción, aprendizaje, uso, navegación, documentación, interface, uso del ordenador, adaptabilidad e innovación. También encontramos numerosas iniciativas encaminadas a la adaptación de las normas ISO para el software y materiales web [2].

Una crítica general a la evaluación de software es la baja fiabilidad y validez de muchos de los instrumentos diseñados para tal fin que circulan entre la comunidad de educadores y formadores, y la necesidad de consensuar estándares de calidad cuya presencia no sea simplemente chequeada mediante listas de control, sino a través de un análisis más cualitativo, contextual e intensivo del material y por parte de distintos agentes evaluadores.

1.2.3 Evaluación de las plataformas tecnológicas

La evaluación de las plataformas tecnológicas está orientada a valorar la calidad del entorno virtual o campus virtual a través del cual se implementa el e-learning. La dimensión y funcionalidad de un campus virtual puede variar sustancialmente según se trate de dar soporte a un curso o cursos o a una institución entera, como es el caso de las universidades virtuales.

En el mercado existen numerosas plataformas estándar con posibilidades de adaptación a las necesidades de los distintos tipos de formación on-line, objetivos de la misma y usuarios, pero también pueden ser creadas ad hoc por la institución que adopta una solución e-learning.

Adquirido o creado, el campus virtual es un elemento dinámico, que evoluciona paralelamente a la propia solución e-learning. De ahí la importancia que adquiere la evaluación formativa o de proceso orientada a la mejora progresiva del entorno virtual. Esta evaluación va ganando terreno a la evaluación diagnóstica, a través de la cual se decide la estrategia más adecuada en función de los objetivos formativos y las necesidades. La competencia del mercado de plataformas virtuales ha obligado a equilibrar los servicios y capacidades de las más importantes, de forma que ya no es tan importante cuál se elija sino como mejorarla una vez implementada. Aunque ello no elimina la necesaria etapa de análisis de necesidades, sin la cual a menudo se realizan grandes inversiones en tecnología sin saber qué se necesitará.

A la hora de determinar la calidad potencial de un campus virtual, se debe poder establecer, de forma general, que sea:

- Estable y fiable
- Tolerante a fallos
- Estándar en implementación de contenidos y recursos tecnológicos
- Ágil y flexible
- Actual e intuitivo para facilitar la inter- acción con el usuario

La mayoría de instrumentos existentes para evaluar campus virtuales (diseñados bien por las propias distribuidoras de plataformas para establecer comparaciones, bien por agencias independientes o universidades), están orientados a determinar las características de los mismos en función de una serie de categorías de análisis (tabla 2).

Sobre la base de estas categorías existen iniciativas que han elaborado modelos o estándares de calidad de plataformas virtuales, entre las que destacamos: el Cybernetic Model for Evaluating of Virtual Learning Enviroments (Britain y Liber, 1999) orientado a la interrelación de los siguientes aspectos: recursos de negociación, coordinación, tutorización, individualización, autoorganización y adaptación; o los Quality Standards on the Virtual Campuses desarrollados por la Universidad Virtual de Pensilvania, (<http://www.vup.org/standards/>), basados en el interface de usuario, los media, el software, los permisos y licencias y la accesibilidad.

Un modelo más cualitativo y flexible es el modelo ACTIONS de Bates (1999), que está especialmente orientado a la toma de decisiones antes de introducir un determinado medio tecnológico, y permite hacer una valoración de las ventajas e inconvenientes con relación a cinco componentes (cuyos términos se corresponden con las siglas del modelo): acces, costs, teaching and learning, interactivity and use facility, organizational issues, novelty and speed. Es la única iniciativa que introduce la estimación de variables económicas, siendo aplicable a cualquier tecnología utilizada por el e-learning.

<i>Categorías</i>	<i>Análisis</i>
-------------------	-----------------

<p>Coste</p> <p>Requerimientos de hardware y software</p> <p>Características</p> <p>Capacidad de desarrollo</p> <p>Herramientas para el estudiante</p> <p>Herramientas para el instructor</p> <p>Herramientas para el administrador</p>	<p>General y de los servicios extras</p> <p>Bajo qué sistemas operativos y navegadores funciona, lenguajes que soporta, ...</p> <p>Técnicas, servicios y soporte que ofrece, ...</p> <p>Posibilidad de implementar nuevas funciones, de realizar reports, tests, soporte de VMRL para crear un entorno tridimensional,...</p> <p>Interacción que permite (síncrona/asíncrona, grupal/individual), acceso a recursos (biblioteca, secretaría, ...), trabajo en grupos, evaluación,... Qué le permiten hacer sin necesidad de programar en html (tests, contenidos, diseño instructivo, ...)</p> <p>Qué le permiten hacer (dar autorizaciones, soporte al usuario, al docente, registro, personalización de mensajes, ...)</p>
---	--

Pese a estas y otras experiencias, coincidimos con la European Network on Intelligent Technologies for Smart Adaptive Systems (<http://www.eunite.org>) en que falta todavía establecer y consensuar criterios a nivel europeo sobre los campus virtuales, relacionados con el establecimiento de estándares de acreditación para su calidad, así como orientaciones en el desarrollo de cursos que se implementan a través de estas plataformas, o el establecimiento de normas relacionadas con el copyright y los derechos de autor, de forma que la evaluación de todos ellos pueda realizarse sobre categorías comparables.

1.2.4. Evaluación financiera

Una de las evaluaciones que está reclamando una mayor atención, especialmente por parte de las empresas, es la relacionada con factores de

medición económicos. La implantación de una solución e-learning requiere una inversión inicial importante, que es injustificable, desde el punto de vista financiero, si finalmente no puede evidenciarse un retorno de dicha inversión. Es lo que en economía se denomina ROI (Return On Investment). Se trata de una fórmula muy simple ($ROI = \text{beneficios} / \text{costes}$) para valorar el retorno esperado de una inversión. Aunque detrás de la simplicidad de conceptos se esconden muchos matices, especialmente cuando se aplican al campo de la formación, y aún más del e-learning.

Cómo medir o determinar los beneficios conseguidos, más allá del valor positivo que por sí misma representa para las personas y la propia organización. El reto no es fácil, ya que los beneficios más productivos de la formación son los más intangibles y difíciles de cuantificar (satisfacción, iniciativa y liderazgo, y las habilidades propias de las personas que configuran la organización), mientras que los más operativos, si bien producen resultados a corto plazo, se deben al resultado de conocimientos mecánicos (aumento de la productividad, ahorro de tiempo) Horton (2001).

Pese a las dificultades evidentes que presenta la evaluación financiera de la formación y del e-learning, existen experiencias y propuestas para su planificación y ejecución, entre las que destacamos un modelo tomado de Duarte (2001b), insertado en el proceso de planificación de la acción formativa. El

modelo presenta una serie ordenada de acciones con una lista de indicadores de estimación para cada una de ellas.

La importancia del ROI como instrumento de evaluación reside en no atribuirle a la formación exclusivamente los beneficios derivados de la mejora de habilidades y conocimientos, sino en poder estimar como esta mejora repercute en los resultados económicos de la organización, convirtiéndose así en un instrumento para controlar la eficacia y eficiencia en la aplicación de las inversiones.

	Acciones		Instrumentos de evaluación
a	Definir los objetivos de la acción formativa	Redactar objetivos acordes con la estrategia de la institución y con la operativa de la unidad de negocio en la que se va a implementar	Frases concretas y concisas que expresen <i>qué</i> resultados esperamos conseguir
	Definir los objetivos de aprendizaje de los participantes	Redactar objetivos de aprendizaje de acuerdo con los participantes	Compromiso de cada uno de los participantes con sus objetivos al inicio de la acción formativa Tabla de relación entre los objetivos de aprendizaje y los resultados esperados
	Determinar la modalidad formativa (presencial, e-learning, dual, etc.)	Analizar qué modalidad formativa puede alcanzar mejor los objetivos esperados de forma más eficiente	Tabla comparativa de ventajas e inconvenientes de las modalidades de formación valoradas Valoración de las resistencias o motivaciones existentes en cada modalidad Valoración de los costes de introducción de una nueva modalidad formativa en función de los resultados esperados

	Determinar los beneficios de la acción formativa	Exponer los beneficios esperados, tanto cuantificables como no, para la organización y para los participantes de la modalidad de formación elegida Establecer los resultados esperados Exponer los resultados en datos cuantificables, señalando también logros y beneficios cualitativos esperados Establecer los niveles de satisfacción esperados (clientes / proveedores; internos o externos)	Lista de beneficios concretos esperados expresados en porcentajes de mejora Lista de beneficios internos a la unidad de negocio, propios de los participantes y de los clientes y proveedores de la unidad Lista de logros expresados en períodos cortos de tiempo Plan de comunicación, interno y/o externo a la propia unidad de negocio, que incluya los objetivos y los logros esperados
	Determinar los costes de la acción formativa	Exponer los costes de la modalidad elegida, tanto para la institución como para la unidad de negocio o los participantes	Lista de costes de la acción formativa
	Calcular el ROI de la acción formativa	Analizar los beneficios, las inversiones y los costes en función de los objetivos y resultados esperados	Cálculo del ROI Publicación del cálculo del ROI
Implementación	Desarrollo de la acción formativa	Desarrollar la acción formativa de acuerdo con la planificación prevista	Recogida de información durante el proceso Información sobre el proceso a los participantes y a los beneficiarios de la acción formativa Corrección de errores de desviación sobre la planificación
Evaluación	Evaluación de la acción formativa	Evaluar la acción a partir de los diferentes instrumentos realizados	Análisis y difusión de los resultados de la evaluación
	Evaluación del diseño realizado e implementado	Valorar mejoras a realizar en el diseño de futuras acciones formativas a partir de las omisiones, debilidades o fortalezas del diseño actual	Lista de elementos a incorporar, mantener o evitar en futuros diseños de acciones formativas

1.2.5 La evaluación de enfoque global

Es la evaluación que tiene presente el conjunto total de elementos que intervienen en una solución e-learning a la hora de establecer líneas y criterios para gestionar o evaluar su calidad.

1.2.6. Evaluación y gestión de la calidad

Sustentada por planteamientos del enfoque sistémico (Pérez Juste et al., 2000) y del paradigma actual de la complejidad (López Rupérez, 1997) y basada en el concepto de mejora continua (Deming, 1981), la gestión de la calidad se distingue por su enfoque global e integral, siendo una estrategia organizativa y una metodología de gestión que hace participar a todos los miembros de una organización con el objeto fundamental de mejorar continuamente su eficacia, eficiencia y funcionalidad.

La implantación de un sistema de calidad en una organización viene fundamentada por los siguientes principios (González, 2000: 66-68):

- Proceso orientado a la satisfacción de las necesidades y expectativas de los destinatarios
- Mejora permanente de todo lo que la organización pretende alcanzar sobre la base de unos objetivos claros y explícitos
- La garantía de la calidad de los procesos internos como medida para alcanzar la calidad del producto
- La prevención en lugar de la supervisión y detección de errores

Además de la importancia de:

- El liderazgo
- El trabajo en equipo, la calidad es tarea de todos
- La resolución sistémica de los problemas
- Basar la toma de decisiones en datos objetivos
- La agilidad en la transmisión de información
- La formación de las personas implicadas

Sobre la base de estos principios se han generado diferentes herramientas, como las normas ISO y los modelos de evaluación propios de la TQM (Total Quality Management).

Una de las herramientas más extendidas es la aplicación de las normas ISO, ese conjunto de normas (de carácter no prescriptivo), que simplemente exigen que una organización: a) defina y planifique sus procesos, b) los documente de manera correcta, c) compruebe su actitud, y

d) garantice el control y revisión de los mismos.

Los principios sobre los que se articula este conjunto de normas son los siguientes:

- La organización posee objetivos de calidad claros
- Existen acuerdos claros entre todos los participantes
- La organización posee los recursos necesarios para conseguir el nivel requerido de calidad
- Todos los procesos y sistemas están sometidos a controles, con evaluaciones y modificaciones cuando sea conveniente
- Todo lo necesario para garantizar la calidad se halla documentado
- Los registros de la calidad permiten una verificación y una comprobación de la garantía de la calidad

La interpretación de estos términos aplicada a contextos de e-learning queda recogida en la tabla 4, a modo de propuesta, y recogiendo las recomendaciones de Van De Berghe, W. (1997)

<i>Término de las ISO</i>	<i>Interpretación para una organización (educativa o formativa) que adopta una solución e-learning</i>
Proveedor	La institución u organización que imparte la enseñanza o formación
Clientes	Alumnado o participantes
Producto	Curso, programa, currículum
Cuadros ejecutivos	Dirección general/Dirección de centro/Rectorado/Comité de dirección
Contrato	Todos los tipos de acuerdos con los clientes: matriculación, acceso al sistema de comunicación, distribución de materiales, tutorización, acreditación,
Diseño	Definición de las especificidades de la plataforma tecnológica y sus capacidades instructivas y administrativas y los sistemas de seguridad
Compras	Adquisición de recursos materiales y humanos y servicios necesarios: contratación de profesorado, licencias de software, compra de hardware, ...
Procesos	Proceso de impartición de la enseñanza o formación: metodología (trabajo en grupo, individual), uso de los espacios de comunicación, materiales (on-line, off-line, multimedia, textuales), evaluación (presencial o a distancia), ...
Inspección y pruebas	Evaluación de la acción formativa, del profesorado, de los materiales, de la tecnología,por parte del alumnado o participantes
Calibración	Validación de los sistemas de evaluación utilizados

Los modelos de la TQM, como EFQM (Euro- pean Foundation Quality Management) o el más recientemente implantado CMI [3] (Cuadro de Mando Integral) poseen unas fases de implementación y una serie de áreas clave o criterios básicos de calidad en torno a los cuales se articula el modelo.

Se trata en todos los casos de herramientas de evaluación de uso frecuente en las organizaciones y en la actualidad se están adaptando a los contextos de formación virtual. Pero como han advertido algunos autores (Mateo, 2000b, Barbe- rá, 2001) se trata de modelos más próximos a la gestión (management) que a los procesos educativos. Estos modelos enfatizan fundamentalmente los aspectos de gestión organizativa, satisfacción del cliente, o relación coste-beneficio, aspectos sin duda importantes, pero insuficientes en una actividad cuyo núcleo es la enseñanza- aprendizaje.

Barberá (2001) propone como dimensiones que deben ser evaluadas en la formación virtual: el escenario en el que se produce la acción educativa (bases psicopedagógicas, estructura general del sistema, etc.); las propuestas de los participantes involucrados en el proceso de instrucción (motivaciones, objetivos y demandas cognitivas);

los agentes instruccionales (roles de estudiantes, de docentes, y de la misma institución); la intervención y la interacción educativa (organización de la actividad educativa, patrones de interacción y discurso virtual); y la construcción misma del conocimiento (características del conocimiento, dinámicas y tipos de construcción).

Muchas de estas dimensiones son recogidas a través de la práctica del benchmarking, que está cobrando una gran importancia en la evaluación del e-learning.

1.2.7. Prácticas de evaluación del e-learning basadas en el benchmarking

El benchmarking es el proceso que permite a un centro u organización compararse con otro que obtiene resultados excelentes de calidad, con el fin de emularlo. En el ámbito del e-learning este sistema pretende ofertar herramientas e indicaciones para mejorar las acciones a partir de la observación, comparación y cooperación basada en las buenas prácticas.

1.2.8 LA CALIDAD DE LOS PROGRAMAS DE EVALUACIÓN Y DE LOS INSTRUMENTOS QUE LOS INTEGRAN

INTRODUCCIÓN

La evaluación de los aprendizajes se realiza sobre la base de un programa que, al servicio de la enseñanza y del aprendizaje, está constituido por un conjunto de instrumentos de evaluación. Dado que cada tipo de instrumento permite evaluar diferentes aspectos de los aprendizajes de los alumnos, es menester garantizar la pertinencia y calidad técnica del programa considerado integralmente como una estructura, así como la de cada uno de sus componentes.

Un principio general que rige el diseño de un programa de evaluación es el de la economía de tiempo. El tiempo destinado a situaciones de evaluación puntual, diferenciadas de las situaciones de enseñanza, debe ser el más breve posible, ya que el factor más crítico de la educación formal es la carencia de tiempo suficiente para enseñar con profundidad toda la información y para desarrollar todas las competencias que los alumnos deben alcanzar.

Las condiciones en que se desenvuelve actualmente la educación escolar en todos los niveles implican, precisamente, la necesidad de sortear algunas dificultades para la evaluación porque las clases suelen ser numerosas como producto de los muy significativos y deseados procesos de democratización de la educación. El número de alumnos ha aumentado en los niveles de educación básica, secundaria y superior, planteando el requerimiento de encontrar las técnicas adecuadas para la enseñanza de grupos numerosos, así como las técnicas de evaluación que puedan acompañar adecuadamente a una enseñanza que promueva aprendizajes significativos en todos los alumnos, que evite el fracaso escolar, la repitencia y la deserción, y que permita hacer efectivo el propósito de igualar las oportunidades educativas reconociendo

los derechos a la igualdad así como a la diversidad de los rasgos personales de los alumnos.

Por ello, la calidad de un programa de evaluación está sujeta, desde un punto de vista pedagógico y por esto, también, ético y político, a su capacidad para evaluar justa y equitativamente a grupos numerosos integrados por alumnos diferentes.

No es posible hablar de la evaluación de los aprendizajes al margen de los procesos de enseñanza y aprendizaje que los han generado. Una condición fundamental para el diseño de programas de evaluación, en cualquiera de los niveles que se deseen o se deban evaluar en la institución escolar, es que tienen que ser consistentes con los proyectos de enseñanza y aprendizaje de la institución ya que sólo alcanzan su propósito de servir como dispositivos para el perfeccionamiento de los resultados de la educación cuando se convierten en juicios de autoevaluación, tanto para los alumnos como para los docentes y las autoridades de la escuela y del sistema.

Para lograr una buena configuración de los procesos de enseñanza, el docente debe tomar algunas decisiones adecuadas. Las decisiones de los docentes están fundadas sobre las concepciones que ellos tienen, entre otras, acerca de qué es enseñar, qué es aprender, cuál es la naturaleza de los conocimientos que los alumnos deben adquirir y qué funciones tiene que cumplir la escuela. Hay, pues, un conjunto de teorías o ideas más o menos explícitas, o a veces, implícitas, basadas en el sentido común o en teorías científicas, según los casos, a través de las cuales cada docente procura construir explicaciones de los procesos y las funciones relacionadas con la intervención pedagógica. Esas

explicaciones le sirven de fundamento cuando diseña su proyecto de enseñanza.

Los docentes también ponen en juego la información que poseen acerca de lo que está ocurriendo con sus alumnos y las características que suponen que éstos tienen como individuos y como grupo de clase, ya que ellos son los destinatarios de la enseñanza. En esta tarea, la referencia se centra en la necesidad de prever cuáles habrán de ser los impactos más duraderos y los efectos más inmediatos que posiblemente tengan sus propias acciones de enseñanza sobre los alumnos. Para efectuar esas previsiones, el docente debe poseer información de distintos tipos. Ella puede ser recogida por medio de la observación del desarrollo de los procesos de enseñanza o puede surgir de las respuestas de los alumnos colocados en situaciones creadas especialmente por los docentes con ese fin.

Si se parte de estas premisas, se advierte que la importancia del papel que desempeñan las teorías explicativas de cómo ocurren estos procesos, qué modalidades adoptan en su presentación y cuáles son las variables que inciden en ellas, obliga al docente a efectuar el análisis y la reconstrucción de dichas explicaciones teóricas, de modo de trabajar con aquellas a las que estima más apropiado adjudicar un mayor valor científico por su actualidad, seriedad y rigor.

Pero, además, en lo que se refiere a la información relativa a los individuos concretos o empíricos que conforman el grupo de clase o de la institución, el docente necesita contar, también, con la mejor información. Para obtenerla

puede recogerla de las situaciones naturales tal como ocurren en la sala de clase o fuera de ella, o bien puede generar intencionalmente situaciones específicas. La evaluación tiene un estado inicial que puede construirse a partir de dos tipos de operación: una modalidad se constituye sobre la base de recoger datos a través de la observación sistemática o asistemática o de algún otro medio no diferenciado del proceso de enseñanza volcados en la elaboración de sus respectivos registros y, la otra, con características bien diferenciadas de la anterior, consiste en obtener datos en situaciones especialmente diseñadas para recoger una información que se considera relevante como indicadora del estado de situación de los aprendizajes que se quieren evaluar.

Además de esta operación inicial, en alguno de sus dos tipos básicos, para que haya una evaluación con sentido es indispensable que existan criterios que permitan construir juicios de valor acerca de lo que la información recogida significa en términos de aprendizaje de los alumnos. La primera operación, por lo tanto, consiste en recoger información; y la segunda, en analizar e interpretar la información recogida sin que, necesariamente, el orden enunciado implique la diferenciación temporal o el establecimiento de una secuencia de estas dos operaciones. La interpretación permite completar el proceso de evaluación y conduce a la construcción de juicios de valor que son característicos de la evaluación como función didáctica. No hay evaluación sin juicios de valor. Debemos mencionar aquí, sin embargo, que algunos autores discrepan con esta afirmación. Jean Cardinet (1989), por ejemplo, advierte, como veremos más adelante, acerca de los riesgos que implica el uso de juicios de valor en la evaluación.

El juicio de valor puede ser construido de dos maneras. Según una de ellas, puede responder a estados afectivos del docente, relacionados con actitudes de aceptación o rechazo, de agrado o desagrado ante determinada conducta del alumno, observada directamente, inferida de conductas observadas o, también, sólo supuesta más o menos indirectamente a partir de la interpretación de variadas situaciones. El juicio de valor puede surgir, igualmente, como producto de una elaboración seria y rigurosa de la información recogida sistemáticamente, a partir ya no de emociones, sentimientos o actitudes, sino de una base de conocimientos que permite fundamentar el juicio de valor de una manera que pretende ser "objetiva". Ante estas dos maneras diferentes de construcción es posible que, por la naturaleza misma de los actos pedagógicos, en el ejercicio integral del papel del docente se requiera poner en juego las dos modalidades de construcción de los juicios de valor, buscando el modo de articularlas adecuadamente, entendiendo que el docente es una persona y que la relación educativa es una relación entre personas.

La evaluación tiene que contar con criterios que permitan analizar e interpretar la información recogida con base en una teoría orientada a responder a la serie de cuestiones que antes mencionamos. Entre ellas incluimos la teoría referida a la evaluación como función didáctica. Si el docente tiene que contar con teorías explicativas acerca de cómo aprenden los alumnos y sobre qué es enseñar, también deberá trabajar con teorías acerca de la evaluación, su relación con los procesos de enseñanza y aprendizaje, el lugar que ocupa la evaluación en la administración de un proyecto de enseñanza, qué alcances y qué límites debe tener.

Es siempre importante insistir en la necesidad de que todas esas teorías sean consistentes entre sí y conformen, en su conjunto, una base de percepción, pensamiento, sentimiento y conocimiento que permita fundar y justificar la acción docente.

Estos principios no suponen, sin embargo, que exista una única concepción científica acerca de las cuestiones que hemos mencionado, incluyendo entre ellas a la propia evaluación. Así como hay diversas concepciones acerca del universo, las hay también sobre la enseñanza y el aprendizaje.

Cuando se adopta una concepción acerca de la enseñanza, de lo que la enseñanza debe ser y no sólo acerca de lo que ella es, esta concepción normativa conduce a una idea normativa de la evaluación, de lo que la evaluación debe ser. Así pues, para una teoría normativa de la enseñanza, no todas las concepciones sobre la evaluación son válidas; algunas son aceptables, otras no lo son.

La evaluación es un recurso indispensable para el perfeccionamiento de los dos procesos que nos interesan, la enseñanza y el aprendizaje. Encarar una acción docente significa actuar para mejorar las probabilidades de que se produzcan procesos deseables de aprendizaje. Pero es menester señalar que, aunque por lo general cuando hay enseñanza se produce algún aprendizaje, éste no siempre coincide con los propósitos del docente. La actitud del docente frente a la evaluación deberá ser abierta, coincidiendo con su teoría de la enseñanza. Si se está de acuerdo con la idea de que, al enseñar, el docente no debe desarrollar una intervención caracterizada por su

unidireccionalidad en la que la única voz a escuchar es la del propio docente sino que hay que dar lugar a la voz del alumno, esto es, a la manifestación de su capacidad para pensar y construir significados, del mismo modo en el proceso de evaluación debe encontrar el alumno un lugar para expresar los significados desde su propia perspectiva.

Cuando se opta, entonces, por concepciones de la enseñanza en las que se manifiesta nuestro respeto por la capacidad de producción personal del alumno, los instrumentos de evaluación que utilicemos deberán ser coherentes con esa postura.

En la actualidad disponemos de una cantidad importante de instrumentos diversos, creados muchos de ellos a lo largo del siglo XX. Consisten, por un lado, en las llamadas "pruebas objetivas", diseñadas a partir del programa de investigación científica del conductismo, con una serie de especificaciones técnicas importantes en cuanto a la construcción de ítems y al análisis e interpretación de resultados. También contamos con un conjunto de normas técnicas para la utilización de las denominadas "pruebas subjetivas"- a veces llamadas "tradicionales"- en las que, a diferencia de las pruebas objetivas que sólo admiten una única respuesta correcta, se aceptan diferentes niveles de calidad de respuesta. En este último campo, las normas técnicas desarrolladas han apuntado a mejorar la construcción, la administración y la interpretación de los resultados, con el propósito de mejorar también la confiabilidad y la validez de las pruebas subjetivas.

En este siglo, el trabajo realizado en el campo de la evaluación didáctica ha sido muy profuso, ya que se ha considerado que se trataba de un tema prioritario en las teorías de la enseñanza. No es extraño, por ejemplo, que para las diversas teorías conductistas del aprendizaje, la evaluación se convirtiera en un aspecto central de su desarrollo teórico y experimental y de su propuesta didáctica, hasta el punto de sostener que se debía limitar lo que se enseñaba a aquello que era posible evaluar en los términos que la teoría aceptaba con carácter exclusivo, es decir, cuantificando los rasgos de las conductas observables. Dada la significativa importancia que se otorgaba a la evaluación de los aprendizajes y a la simplicidad técnica del problema planteado, se comprende la gran fertilidad de las teorías conductistas en el dominio teórico y práctico de la evaluación.

Consideramos que se trata de una función fácil de resolver técnicamente en el marco de las teorías del conductismo porque se limita a evaluar lo que se puede observar. Las corrientes cognitivas, en cambio, se proponen evaluar los procesos mentales. Su problema es más difícil de resolver que el que se planteaban los conductistas y sus respuestas técnicas son más dificultosas de construir. La cuestión de la evaluación, en cambio, es menos importante para las didácticas humanistas que plantean una relación diferente entre docente y alumno en la que los aprendizajes son aún más difíciles de evaluar que en las propuestas didácticas cognitivas.

Debemos señalar, empero, que el programa conductista caracterizado por su gran capacidad de producción de instrumentos técnicos de evaluación durante varias décadas del siglo XX no fue el único programa que realizó trabajos en este tema; no fue ni siquiera el que ejerció una hegemonía

completa en la evaluación. En el campo teórico y práctico, simultáneamente con el trabajo que en este ámbito fuera realizado por los psicólogos conductistas, también se trabajó con otros instrumentos de evaluación que respondían a perspectivas diferentes. Algunas corrientes europeas que desarrollaron trabajos importantes en este terreno lo hicieron desde enfoques diferentes, alternativos a las producciones conductistas. Entre ellos, por ejemplo, se destacan los que entregan al alumno la responsabilidad de seleccionar el material, organizarlo y encontrar formas adecuadas de expresar su producción personal, cuestiones que no pueden ser evaluadas por medio de la administración de pruebas objetivas. De este modo, y aún en el ámbito de mayor influencia de las teorías conductistas, se emplearon siempre y se enseñó a los docentes a construir, analizar e interpretar los resultados de las llamadas "pruebas subjetivas" o "tradicionales".

Es importante hacer notar que la gran disponibilidad de instrumentos de diferente carácter, alcance y función, es un factor de enriquecimiento del abanico de posibilidades que se abren para el diseño de programas de evaluación. El problema que se ha de resolver es el de su adecuada selección y organización de modo de alcanzar consistencia con el proyecto docente.

Uno de los productos más importantes del trabajo de los diversos programas de investigación sobre la evaluación en este siglo ha sido, precisamente, el mejoramiento de la normativa de construcción de instrumentos y análisis de resultados, lo cual nos permite abordar el problema de la evaluación contando con una amplia variedad de instrumentos que incluyen pruebas escritas y orales y pruebas de observación de procesos y de producciones, utilizables de diversas maneras según las necesidades de la enseñanza. El

aporte del conocimiento desarrollado en las últimas décadas sobre los procesos cognitivos da un nuevo sustento a las teorías y técnicas evaluativas y permite confirmar o modificar las prácticas instaladas en la educación formal desde hace largo tiempo en algunos casos, y en otros recientemente, como resultado de la introducción de nuevas propuestas que es necesario reanalizar a la luz de la investigación actual.

LOS PROGRAMAS DE EVALUACIÓN

Si nos preocupamos hoy por determinar cuáles son los ejes en torno de los cuales gira esta cuestión, tendríamos que comprender, en primer lugar, que el propósito principal de la evaluación no se puede lograr si la evaluación no se convierte en autoevaluación tanto para el docente cuanto para el alumno. La evaluación, en segundo lugar, debe ser consistente con las concepciones de la enseñanza y del aprendizaje. En tercer lugar, la resolución técnica que se le dé debe permitir evaluar todos los aspectos que están comprometidos con los procesos de aprendizaje, lo cual supone desarrollar programas complejos de evaluación en los que se empleen una cantidad de instrumentos diversos y donde cada técnica sea adecuada para evaluar diferentes aspectos.

Los programas de evaluación deben ser diseñados por los docentes, como parte de la programación didáctica de sus cursos. Para ello, el docente debe conocer en profundidad la o las teorías de la evaluación y la variedad de instrumentos que existen. Estos conocimientos le permitirán diseñar programas coherentes con la programación de la enseñanza combinando instrumentos diversos, así como crear nuevos instrumentos de evaluación acordes con la multiplicidad y heterogeneidad de los propósitos de su enseñanza. Es también indispensable, si quiere desarrollar un buen programa

de evaluación, que el docente conozca las normas técnicas para la construcción, administración, análisis e interpretación de resultados.

Queremos señalar que, en la medida en que un docente es responsable de la elaboración de su programación didáctica, deberá tener la misma libertad para diseñar y administrar su programa de evaluación. Los grados de libertad en la toma de decisiones deben ser equivalentes. Si así no ocurriera y si el docente tuviera libertad para decidir algunos aspectos de su proyecto de enseñanza y no la tuviera para decidir acerca de su programa de evaluación porque ésta le es impuesta desde el exterior, la evaluación no cumpliría otra función que la de servir de control, resultando minimizados sus posibles efectos positivos de mejoramiento de los procesos y de los resultados de esos procesos de enseñanza y de aprendizaje. Hay que recordar siempre que la libertad de quien enseña debe tener su correspondencia en la libertad de quien evalúa y en la libertad de quien aprende. Y, respecto de este último es importante recordar, como lo hace E. R. House (1994) citando a John Rawls, que ". . . respetar al otro como persona moral supone tratar de comprender sus aspiraciones e intereses desde su punto de vista y presentarle consideraciones que le hagan posible aceptar las limitaciones de su conducta" (pág. 129).

LOS DETERMINANTES DE LA CALIDAD DE LA EVALUACIÓN

Un programa de evaluación y cada uno de los instrumentos de evaluación que lo integran deben cumplir con ciertas condiciones. Mencionaremos aquí algunas de ellas: debe permitir evaluar los aprendizajes que la institución y el o los docentes esperan que los alumnos desarrollen. Las consignas de trabajo que integran el instrumento deben ser seleccionadas, en consecuencia, de modo de permitir que tanto la presencia de una respuesta como su ausencia

sean significativas en relación con los aprendizajes que se ha buscado promover. En términos generales, las consignas no deben producir respuestas que resulten irrelevantes para la evaluación de los aprendizajes previstos. Esto no significa, sin embargo, que se descarten las respuestas irrelevantes de los alumnos cuando aparecen, ya que a través de esos resultados en apariencia disfuncionales, a veces incomprensibles, el docente puede advertir errores que sí pueden ser altamente significativos y que requieren especial atención. El instrumento, aunque siempre haga una lectura parcial de lo aprendido por los alumnos, debe proponerse registrar no sólo los éxitos y fracasos sino también el origen de esos fracasos. El instrumento debe presentar el grado de organización suficiente para que la apreciación que efectúa del aprendizaje permita desprender algunas conclusiones acerca del desempeño presente y futuro del alumno, en cuestiones específicas pero también con visión integral.

Cada uno de los instrumentos de evaluación resuelve estos problemas de diversas maneras. Actúa como un reflector que en el escenario ilumina a algunos personajes y deja en la penumbra a otros, que sin embargo están allí, presentes, pero con una clase diferente de presencia. Por esta razón, la elección de los instrumentos de evaluación adecuados a la hora de diseñar el programa de evaluación de un curso, por ejemplo, constituye una de las decisiones más importantes para garantizar el valor didáctico de la evaluación en relación con los procesos de enseñanza y de aprendizaje.

Existe una variedad de modelos de instrumentos de evaluación. Cada uno de ellos, tomado aisladamente, es insuficiente para obtener información sobre el aprendizaje de los alumnos. El diseño del programa exige, pues, la combinación de instrumentos diversos para obtener una cobertura adecuada.

La eficacia de la evaluación depende, entonces, de la pertinencia de la combinación de diferentes instrumentos, de la oportunidad en que se administran y de la inteligencia y propiedad del análisis e interpretación de sus resultados.

Veamos, entonces, los criterios o requisitos que deben tenerse en cuenta en el momento de tomar decisiones acerca del programa de evaluación de un curso, por ejemplo, y sobre cada uno de los instrumentos de evaluación que lo integran. Las características generales que los instrumentos de evaluación deben reunir, aunque con distinto grado de relevancia, son cuatro: validez, confiabilidad, practicidad y utilidad.

VALIDEZ

Se dice que un instrumento de evaluación es válido cuando evalúa lo que se pretende evaluar con él. Como un instrumento es utilizado para apreciar ciertos logros de aprendizaje de un cierto grupo de alumnos en una cierta circunstancia y en un cierto momento de su proceso de aprendizaje (al inicio, en el transcurso o al final), la validez de un instrumento no puede ser determinada de manera absoluta, sino siempre en relación con su adecuación a los propósitos y situación específica de su aplicación. Como lo señala Henry E. Garrett (1966), "un test es válido para un propósito particular o en una situación especial; no es universalmente válido". Cuando se quiere determinar si un instrumento es válido se requiere, entonces, información acerca de los criterios que han presidido su construcción y su administración. Los criterios son, entonces, externos a la evaluación misma. Esta cuestión sería relativamente simple de resolver si todos los propósitos de la enseñanza, en todos los niveles de decisión y para todos los actores, fueran coincidentes.

Pero rara vez ocurre esto. A veces los objetivos del docente coinciden plenamente con los de la escuela y los de ésta con los de la jurisdicción y los de esta última con los de la nación. Pero la plena armonía de todos esos niveles parece más improbable que posible de alcanzar y, más aún, ni siquiera se puede afirmar que, en lo pormenorizado de los aprendizajes escolares, sea absoluta y efectivamente deseable. Las diferencias en la educación surgen de la acción y la resolución de cada docente y de cada alumno. Se podrán encontrar, en consecuencia, criterios generales, pero también criterios diferentes. Y no tiene sentido hablar de la validez de un instrumento si no se ha validado el criterio. Para lo cual se requiere, a su vez, una validación mediante criterios más amplios. Según Robert L. Ebel (1963), "esto requiere un criterio para validar el otro criterio y así sucesivamente ad infinitum.

En la Didáctica, ese proceso de validación al infinito podría ser reemplazado por una teoría de bases firmes, tanto desde un punto de vista científico como filosófico. De otra manera, y ante la obligación ineludible de evaluar a sus alumnos, los docentes se encontrarían en la situación de ser incorporados como nuevos personajes, junto a Aquiles, la tortuga y la flecha, de las aporías de Zenón de Elea.

No se trata, en consecuencia, de validar instrumentos y criterios como en un juego de permanente retroceso hacia los principios, sino de apoyarse en principios didácticos fuertes, con cierto grado de generalidad y que permitan fundamentar racionalmente las decisiones de diseño del programa y de los instrumentos. Veámoslo en algunos ejemplos.

Si para un docente el propósito principal de su enseñanza ha sido procurar que sus alumnos aprendan a resolver un cierto tipo de problemas, el instrumento de evaluación válido será aquel que introduzca problemas al alumno para que éste los resuelva. Deberá tener presente, sin embargo, que no obtendrá la misma información acerca de lo que sus alumnos saben cuándo en el instrumento se plantean problemas que corresponden sólo al tipo de problema cuya solución los alumnos deben hallar, que si también incluyen otros problemas de tipos diferentes, de modo que los estudiantes deban reconocer a qué tipo de problema corresponde cada uno de

los planteados por el docente. Se podría afirmar que, en tanto el propósito didáctico no habría de ser que el alumno resuelva sólo una clase de problemas, la capacidad de discriminar clases de problemas debería estar presente en la formulación de los propósitos de un programa de enseñanza, por lo que esa segunda alternativa, que incluye diversos tipos de problemas en el instrumento, es más válida que la primera, que presenta solamente un único tipo de problema. Así, pues, para lograr que la construcción del programa o de un instrumento de evaluación sean válidos, es menester poner en juego no sólo una lectura fragmentada de los propósitos de la enseñanza sino principios didácticos más generales.

En otro sentido y refiriéndonos a otro ejemplo, veamos el papel que tiene el tipo de instrumento: si se trata de construir un examen oral o un coloquio válidos, hay que definir previamente preguntas centrales. Éstas deben referirse a cuestiones fundamentales del dominio de conocimiento que el alumno debe aprender. Pero la validez no llega a estar asegurada de este modo ya que en estos tipos de instrumentos de evaluación, el examen oral y el coloquio, la

característica que les es propia y que hay que desarrollar adecuadamente es la elaboración de preguntas o comentarios del docente a partir de las respuestas del alumno. Si ellas no le permiten al alumno corregir sus respuestas anteriores, por ejemplo, o no lo animan a profundizar respuestas que inicialmente eran superficiales, o si no lo inducen a la elaboración de asociaciones y conclusiones reuniendo respuestas fragmentarias, la técnica del examen oral o del coloquio están mal empleadas y podrían ser reemplazadas por un instrumento de respuesta escrita. Como se ve, la validez de los instrumentos también depende de su apropiada construcción en relación con las características específicas que los definen.

Se puede observar, a través de lo expuesto hasta aquí, que la validez nunca es absoluta. En primer lugar, porque siempre debe ser referida a un criterio externo. En segundo lugar, porque sólo se puede hablar de validez mayor o menor, ya que un instrumento nunca es "completamente" válido pues sólo permite hacer inferencias acerca de los saberes de los alumnos. Y finalmente, porque la validez está condicionada a un ajuste que nunca puede ser perfecto entre el saber aprendido que se quiere evaluar y los medios técnicos, esto es, los instrumentos disponibles para evaluarlo. Con frecuencia se obtiene mayor validez a través de la combinación feliz de distintos tipos de instrumentos.

Se diferencian diversas clases de validez. Desarrollaremos sintéticamente sus principales modalidades.

Validez de contenido

Se dice que un instrumento tiene validez de contenido cuando representa una muestra significativa del universo de contenido cubierto por el curso o la unidad didáctica, la clase o una serie de clases. La validez de contenido, por lo tanto, requiere que la selección de contenidos que se ha efectuado para construir el instrumento sea equivalente a la de cualquier otro conjunto de contenidos que se hubiera podido seleccionar con ese fin. Se trata, por supuesto, en todos los casos de contenidos importantes por su significación en el aprendizaje de los alumnos.

La dificultad que la construcción de un instrumento debe superar es la necesaria distancia que existe entre la gran cantidad y variedad de los contenidos de toda enseñanza y la brevedad exigida también para los tiempos destinados a una evaluación de los aprendizajes, si no se quiere fatigar a los alumnos, lo cual conduciría a desvirtuar el significado cognitivo de sus respuestas.

Esa distancia obliga al docente a elegir los temas, informaciones, competencias, habilidades, destrezas, procedimientos u otros que incluirá en el instrumento de evaluación. Cuanto mayor es la diferencia entre tiempos destinados a la enseñanza y al estudio y tiempos destinados a la evaluación, más difícil es lograr una validez de contenido aceptable, es decir, que permita asegurar que los resultados obtenidos por el alumno en una instancia de evaluación serían semejantes a los que obtendría en otra distinta. La justicia de la evaluación se pone en juego, entonces, cuando se efectúa el muestreo de los contenidos que se incluirán en el instrumento a través de las consignas de trabajo (por ejemplo, las preguntas) que éste presenta.

La validez de contenido es también llamada "validez curricular" porque el criterio externo que se emplea para garantizar esta clase de validez es el diseño curricular, ya que en éste se expresan los propósitos y los niveles de aprendizaje, las secuencias y las sugerencias sobre estrategias e incluso, en ocasiones, sobre actividades de aprendizaje y modalidades de evaluación. Con la libertad que debe caracterizar al manejo del diseño curricular por parte del docente, es indudable que este diseño es un recurso fundamental sobre el que se apoyará la estimación acerca de la validez de contenido de un instrumento de evaluación, así como la fuente en la que el docente encontrará la principal inspiración para realizar una selección adecuada y representativa de los contenidos de su curso. Es preciso recordar que cuando se habla de "contenido" se hace referencia a un concepto amplio que comprende todo lo que se quiere enseñar. Se incluyen, por tanto, informaciones, conceptos, principios, procedimientos, valores, actitudes, hábitos, destrezas motrices y todo aquello que se tiene el propósito de que los alumnos aprendan.

En virtud de que lo que se enseña varía de clase en clase, de docente a docente y de región a región, la validez de los instrumentos también es limitada en relación con sus alcances.

Validez predictiva

Esta clase de validez se refiere a la correlación existente entre los resultados obtenidos en una o varias pruebas combinadas y el desempeño posterior del alumno en aquellos aspectos que corresponden al área evaluada por esas pruebas. Los instrumentos de evaluación que se emplean en la escuela deben tener validez predictiva escolar y extraescolar y deben contribuir a

pronosticar los éxitos y dificultades de los sujetos del aprendizaje tanto en los aprendizajes posteriores de las asignaturas escolares cuanto en sus desempeños ocupacionales o en su actuación social en general.

Un nivel significativo en la validez predictiva es una condición para un buen instrumento de evaluación. Y esto es así por cuanto no hay aprendizaje escolar cuyo único fin consista en procurar una única ejecución exitosa en un espacio escolar y un tiempo muy limitada. El mismo fin de la educación se define por la voluntad expresa de quien enseña de que los aprendizajes trasciendan el espacio y el tiempo escolares. Es el destino mismo de la educación buscar la perdurabilidad de los aprendizajes. De ahí que, aunque haya algunos instrumentos cuyo propósito específico es la predicción - como es el caso de las pruebas de admisión-, todo instrumento de evaluación empleado con fines pedagógicos tenga, entre otras, la función de predecir. Como antes dijimos, la validez de un instrumento no es absoluta, siempre es limitada. Pero el seguimiento de los alumnos para comprobar la capacidad del instrumento para efectuar predicciones adecuadas es una operación de medición relativamente sencilla, al alcance de todos los docentes que deben encararla como parte constituyente de sus programas de evaluación. Comparar los resultados obtenidos por cada alumno en pruebas diferentes, separadas en el tiempo y en las que el docente asume que hay relación entre los contenidos, es una tarea interesante a la hora de evaluar la enseñanza impartida. Si en una prueba se han detectado errores en las repuestas y se ha brindado atención especial a los alumnos que tenían esas dificultades específicas, el seguimiento de los resultados en una prueba posterior demostrará, a la vez, la validez de la prueba anterior y la adecuación del tratamiento didáctico de los errores que se ha encarado en la acción de enseñanza.

Debemos señalar, sin embargo -aunque esto no exima al docente de hacer el esfuerzo de evaluar la validez predictiva de los instrumentos que utiliza- que los aprendizajes que la escuela promueve tienen plazos más o menos prolongados para su logro pleno. Se han de diferenciar aquellos logros que el alumno puede demostrar en un plazo muy corto, casi contemporáneo con el período de su enseñanza en la escuela, con los aprendizajes que requieren un período algo más prolongado de procesamiento, trabajo y elaboración para una apropiación profunda y aquellos otros aprendizajes que, por el alcance temporal que tiene la maduración personal que requieren, no podrían ser evaluados durante la permanencia del alumno en la institución educativa y ni siquiera en un lapso próximo a su egreso. En estos últimos casos la apreciación de la validez predictiva de los instrumentos se ve limitada por razones técnicas difíciles de superar por el docente y por la institución. Queda en pie, empero, la necesidad de poner atención constante en la capacidad de predicción que tienen los instrumentos de evaluación que se emplean en la educación formal.

Validez de construcción

El programa de evaluación y cada uno de los instrumentos que lo integran debe estar construido de acuerdo con los principios de la o las teorías didácticas que sostienen el proyecto pedagógico. Debe haber coherencia entre las teorías y las estrategias de enseñanza que se han de emplear o que se han utilizado, las teorías generales y específicas que describen y explican los procesos de aprendizaje que el programa de enseñanza se propone que los alumnos desarrollen y las teorías que fundamentan las elecciones de los tipos de instrumento, la secuencia empleada, la articulación de la enseñanza y de las instancias de evaluación que conforman el programa de evaluación diseñado. Lo mismo ocurre con las normas técnicas que deben respetarse

cuando se construye cada tipo de instrumento de evaluación. De la coherencia de las teorías y el buen uso de las normas técnicas depende la validez de construcción del programa y de los instrumentos de e valuación.

Es éste un aspecto que debe ser particularmente cuidado por los docentes porque si la evaluación no tiene un fundamento teórico válido, las decisiones que se tomen a partir de la información recogida por esa vía y de las interpretaciones que se hagan de ella carecerán de legitimidad.

Puede comprenderse, por lo tanto, la importancia de esta clase de validez ya que ella condiciona las otras que antes hemos mencionado.

Validez de convergencia

Esta modalidad de la validez concierne a la relación que existe entre un programa de evaluación o un instrumento y otros programas o instrumentos de validez ya conocida. Cuando un programa o una prueba de evaluación son reemplazados por otros, por ejemplo, debe demostrarse que son, por lo menos, tan buenos o mejores que los empleados anteriormente. La validez de convergencia puede tener gran importancia, pues, cuando se modifican prácticas establecidas en la evaluación para mejorar su validez, su confiabilidad o su practicidad, o cuando se busca mejorarla ampliando el campo de los aspectos evaluados.

Tomemos, por ejemplo, una situación en la que es necesario encontrar una nueva modalidad de evaluación que permita trabajar adecuadamente con un número muy superior de alumnos al que se acostumbraba atender en la institución. En ese caso, se trata de encontrar un programa e instrumentos más prácticos pero igualmente válidos que los que se empleaban con

anterioridad. Los resultados obtenidos con los nuevos programas deberán ser contrastados, en consecuencia, con los obtenidos por medio de la aplicación de los usados antes. La validez de convergencia se apreciará analizando las coincidencias entre ambas series de resultados de la evaluación. Pero si, en cambio, se procura obtener mayor información acerca de un aspecto significativo del aprendizaje, profundizando el examen de los logros obtenidos por los alumnos en competencias o ejecuciones específicas que hasta ese momento se evaluaban de manera más general, en situaciones diferentes o en tipos de expresión distintos, la elección de los nuevos instrumentos deberá ser sometida a un análisis de convergencia para establecer si los resultados de unas y otras series son coincidentes o contradictorios. Si ocurriera esto último, habría que reexaminar tanto los instrumentos empleados previamente como los nuevos, antes de adoptarlos y dar estabilidad al nuevo programa de evaluación.

Validez manifiesta

Esta modalidad de la validez no se refiere a la validez intrínseca del programa o del conjunto de las pruebas o de cada prueba, sino al modo en que ellos se presentan, aparecen, ante el público.

Su significatividad deriva de que los programas y los instrumentos de evaluación deben ser percibidos por los alumnos, sus padres y el público como instancias adecuadas para la evaluación. No pueden ser vistos como absurdas carreras de obstáculos ni como simples o complejos filtros que tienen el único fin de producir una selección, más o menos azarosa y más o menos injusta, entre los alumnos. La razonabilidad de los instrumentos debe ser visible y explicable. Los instrumentos deben ser vistos como formas

capaces de dar cuenta de los aprendizajes necesarios de los alumnos, tanto en lo que respecta a la forma como a los contenidos evaluados.

La validez manifiesta exige, entonces, que los instrumentos evalúen de manera verosímil los aprendizajes. Sin esto la evaluación pierde transparencia y no puede convertirse en autoevaluación. Al perderse la confianza en ella, pierde de este modo su sentido pedagógico y se requiere efectuar las correcciones y operaciones que resulten necesarias para lograr restablecerlo.

Validez de significado

Dietel, Herman y Knuth (1991) añaden a los distintos tipos de validez que se describen habitualmente, la validez de significado, que se mide en respuesta a la siguiente pregunta: ¿las tareas de evaluación tienen significado para los estudiantes y los motivan para alcanzar su mejor rendimiento?

Semejante a la modalidad de validez manifiesta que hemos descrito, la validez de significado se centra, en particular, en la relación que se establece desde la perspectiva de los estudiantes, entre el programa y los instrumentos de evaluación, por un lado, y los procesos de enseñanza y de aprendizaje, por el otro. La evaluación tiene diversos efectos sobre el aprendizaje. Entre los efectos deseados se encuentra el de proveer información al alumno sobre los progresos y los obstáculos que encuentra en el proceso de su aprendizaje y sobre los errores que debe superar, dándole indicios claros acerca de cuál puede ser el origen de sus dificultades. Pero la evaluación también tiene el efecto de incentivar la motivación del alumno por aprender. La superación de dificultades constituye un aliciente para el aprendizaje, así como la reiteración de los fracasos disminuye el interés, si se perciben las dificultades como insuperables. En ese espacio de dimensiones variables para cada alumno, y que se establece entre los problemas demasiado difíciles para ser resueltos y

los problemas demasiado fáciles en los que no se puede fallar, se encuentra el punto adecuado en el que debe ubicarse la dificultad que presenta la prueba de evaluación para que tenga significado para el estudiante y constituya un desafío en el que éste ponga su mejor esfuerzo de producción.

Validez de retroacción

La evaluación tiene un efecto de retroacción sobre la enseñanza. Un caso clásico es el que se presenta en las evaluaciones externas, que ejercen una influencia normativa sobre lo que se enseña. Conocidos los contenidos y las técnicas empleadas en las evaluaciones, ellas se convierten para los docentes en los modelos de lo que se debe enseñar y aprender. Lo mismo ocurre, con frecuencia, con los aprendizajes encarados por los alumnos, que se conforman a las pautas establecidas por los docentes en las evaluaciones que administran. Es un efecto, en este caso no deseado, de la evaluación, por cuanto, en vez de ocupar el lugar de un apoyo para el mejoramiento del aprendizaje y de la enseñanza, la evaluación se convierte en la verdadera reguladora de esos procesos. Cuando esto ocurre, el peligro es grande, ya que, como dijimos antes, el dominio de los aprendizajes en la educación es mucho más amplio que el que se puede cubrir mediante la evaluación.

La validez de retroacción, que debe ser medida al igual que los otros tipos de validez, impone la realización de acciones de corrección con el fin de restablecer el verdadero sentido didáctico de los procesos de evaluación.

CONFIABILIDAD

Se dice que un instrumento de evaluación es confiable cuando una exactitud en la medición y sensibilidad para la apreciación de la presencia y las diferencias de magnitud de los rasgos que mide.

Para que los resultados obtenidos mediante la administración de un instrumento de evaluación puedan ser considerados dignos de confianza, deben ser estables, permaneciendo semejantes en todas las ocasiones en que se administre ese instrumento u otro similar. Si, por ejemplo, se administra una prueba o se emplea algún otro tipo de instrumento de evaluación y se obtienen ciertos resultados, el instrumento es confiable si, unas semanas después, administrado nuevamente se obtienen resultados similares. Esa constancia de los resultados demostraría, primero, que el papel del azar es muy pequeño y no distorsiona de modo significativo los resultados obtenidos. Y, además, que la administración del instrumento ha podido dejar de lado la influencia de factores transitorios que no deberían tener relevancia en la consideración de los resultados del aprendizaje que se quieren evaluar.

Un instrumento confiable permite aislar los aspectos que mide de otros que para el caso se consideren irrelevantes. Así, por ejemplo, una lámpara colgante puede servir para detectar un sismo cercano, pero no puede ser considerada un instrumento confiable pues, al no encontrarse aislada de otros factores, el viento o una mano, por ejemplo, podría ser accionada por otras fuerzas que falsearían la interpretación de su movimiento pendular. De la misma manera, un instrumento utilizado en la escuela con fines de evaluación de los aprendizajes alcanzados por los alumnos debe permitir la consideración de estos aprendizajes separándolos de otros factores que pueden incidir en

las respuestas de los alumnos, tales como el azar, la fatiga o los estados de tensión nerviosa de los alumnos o de los docentes.

La confiabilidad depende, también, de la exactitud y precisión con que mide el instrumento. En primer lugar, debe detectar si el aprendizaje o el rasgo que es evaluado está presente o ausente (¿saben o no saben x?, ¿tienen y destreza o no la tienen?,

¿Quieren z o no la quieren?, etcétera). Pero debe hacerlo con sensibilidad y sin ambigüedad ante las variaciones de intensidad, profundidad y calidad de las conductas observadas o inferidas.

La confiabilidad de un instrumento depende, pues, de su estabilidad, su exactitud y su sensibilidad. Agregamos a estas condiciones la objetividad, que supone que la evaluación es independiente de la persona que evalúa, esto es, que los resultados reciben una interpretación similar tanto de diferentes docentes como de un mismo docente en distintos momentos.

Entre los factores que afectan la confiabilidad se encuentran los siguientes: la longitud y duración de la prueba, ya que cuando es excesivamente breve su capacidad de cobertura de contenido es escasa y puede estar muy distorsionada por factores de azar. "La muestra limitada de conducta puede dar como resultado una puntuación inestable y poco confiable" (Thorndike y Hagen, 1989). Pero cuando la prueba es excesivamente larga, la duración prolongada puede proporcionar información afectada por la fatiga del alumno. Sin embargo, señala Gronlund (1973) que es preferible el riesgo de diseñar pruebas más largas para que tengan mayor validez de contenido, siempre que

se añadan contenidos diferentes y que no se superpongan con los que ya presenta la prueba. Cuando no se trata de instrumentos para medir velocidad sino de pruebas de competencia o capacidad, el tiempo debe ser suficiente para la realización del trabajo, incluso para estudiantes lentos.

Otro factor que incide sobre la confiabilidad son las condiciones de administración del instrumento. En situaciones puntuales de evaluación formal, la tensión de los alumnos suele ser mayor que en las informales, y la manera en que son tratados por el examinador puede favorecer o perjudicar su desempeño. También tienen influencia sobre los rendimientos las condiciones materiales y físicas de administración de las evaluaciones tales como iluminación, silencio, mobiliario, equipamiento adecuado y comodidad en general de alumnos y docentes.

Alcances de la confiabilidad

La confiabilidad de un instrumento debe permitir cubrir apropiadamente las necesidades planteadas por el programa de evaluación. No todos los instrumentos tienen la misma exactitud y sensibilidad, ya que ella depende de la forma técnica que adoptan y de las características propias que les otorga su construcción específica. Así, en otro orden de mediciones, por ejemplo, un barómetro doméstico puede ser aceptablemente confiable porque las lecturas que permite efectuar brindan la información que, con márgenes de error adecuados a ese uso doméstico, le permiten cumplir con la función a la que está destinado. En una estación meteorológica, en cambio, esas mediciones serían consideradas insuficientemente confiables, pues la sensibilidad y exactitud requeridas para cumplir su función son mucho mayores. Los instrumentos de medición de precisión son los que tienen mayor estabilidad,

exactitud, sensibilidad y objetividad. Pero se debe recordar siempre que todas las mediciones no son absolutas, sino que son relativas a los instrumentos de medición que se emplean.

En cada caso, en consecuencia, habrá de determinarse cuáles son los grados de confiabilidad que resultan necesarios para cumplir con los propósitos que el programa de evaluación asigna a cada instrumento.

No se puede confundir confiabilidad con validez. Es frecuente que los autores de libros sobre este tema afirmen que la confiabilidad es condición necesaria, aunque no suficiente, de la validez (Karmel, 1986; Gronlund, 1973; Popham, 1983; Ebel, 1963; Sachs Adams, 1970; Pidgeon y Yates, 1979). Veamos algunas de estas afirmaciones.

Popham (1983): "Un test carente de (con)fiabilidad no puede ser válido para fin alguno". Ebel (1963): "Para que una prueba sea válida debe ser confiable". Pidgeon y Yates (1979): "Un examen no puede ser válido a menos que sea (con)fiable; pero puede ser (con)fiable sin ser válido". Karmel (1986): "Recuerde que ningún test será válido a menos que sea confiable. Por otra parte, un test puede ser confiable y carecer de validez". Esta exigencia de confiabilidad como condición para hacer posible la validez ha tenido mucha influencia en la teoría y la práctica de la evaluación. Consideramos, sin embargo, que es un principio que debe ser examinado y modificado.

La validez y la confiabilidad de un instrumento de evaluación nunca son absolutas. Hemos visto que para la validez los criterios son los que sirven de pauta y, del mismo modo, que la confiabilidad presenta márgenes de aceptabilidad según el uso que se le quiera dar al instrumento.

Pero por otra parte, en los instrumentos de evaluación de los aprendizajes, cuando se exige confiabilidad como primera condición para lograr cumplir con la exigencia de validez del instrumento, se privilegia un cierto tipo de instrumentos en los que el peso de la subjetividad desaparece. Esos instrumentos de alta confiabilidad son aptos para evaluar un cierto tipo de aprendizajes o rasgos pero no para evaluar todos. Las llamadas "pruebas subjetivas" o "tradicionales" (exámenes o cuestionarios orales, pruebas escritas de ensayo o disertación o de respuesta extensa; pruebas de respuesta restringida o de cuestionario; monografías, tesinas y tesis; evaluación de portafolio; estudios de casos, entre otras) están destinadas a evaluar la producción del alumno sin pautas totalmente estructuradas ya que le dan libertad para escoger la información, organizarla, expresarla y presentarla. En estos casos la validez aumenta a expensas de la confiabilidad. Esta última no es, entonces, condición necesaria de la validez sino, por el contrario, una limitación de la validez en ciertas circunstancias, aquellas en las que se busca que la respuesta del alumno sea original y permita demostrar competencias que no pueden ser evaluadas por medio de instrumentos cuyas respuestas están previamente estructuradas por los evaluadores.

Conviene señalar, sin embargo, que según se trate de propósitos de alcance e impacto muy importante en las vidas de las personas (exámenes de admisión o de selección, por ejemplo) o de las instituciones, la transparencia de las

situaciones de evaluación constituye una exigencia ineludible para la evaluación y que, para ello, se requiere una alta confiabilidad además de validez predictiva.

La confiabilidad es, por lo tanto, una característica importante, difícil de aplicar adecuadamente y que exige un manejo técnico muy correcto de las técnicas de la evaluación puestas al servicio de los propósitos que ésta tiene asignados en los procesos de enseñanza.

PRACTICIDAD

Además de las características mencionadas -validez y confiabilidad- hay que considerar una tercera, de mucha importancia en los programas de evaluación.

La practicidad de un programa o un instrumento resulta de la conjunción de tres aspectos: su administrabilidad, la facilidad de análisis e interpretación de sus resultados y elaboración de conclusiones y, por último, de la evaluación de la economía de tiempo, esfuerzo y costo de su utilización.

Administrabilidad

La administrabilidad de un programa o un instrumento es un producto de varias características:

- I. el tiempo de trabajo de los docentes que insume su diseño y construcción;

2. el tiempo que exige su puesta en práctica;
3. la claridad de la/las consigna/s, es decir, de las directivas que se dan a los estudiantes;
4. los materiales, equipos y lugares especiales necesarios para su administración;
5. la cantidad y preparación de las personas indispensables para la administración, cómputo, análisis e interpretación de los resultados y elaboración de las evaluaciones.

Cuanto menor sea el tiempo necesario para el diseño y la elaboración de las conclusiones y menos especializado el personal necesario para la puesta en práctica y el cómputo, mayor será la administrabilidad del programa o instrumento.

ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS Y ELABORACIÓN DE CONCLUSIONES

Se puede advertir que hay una correlación inversa entre el tiempo que demanda la construcción de un instrumento y el que exige el análisis e interpretación de sus resultados. Es necesario señalar, sin embargo, que el diseño de programas e instrumentos de evaluación siempre requiere mucho tiempo, también cuando se trata de pruebas "subjetivas". El tiempo de "corrección" de las pruebas o de registro de observaciones suele ser prolongado pero la tarea puede ser facilitada por la construcción de claves o grillas de corrección elaboradas previamente a la administración de la prueba, sea escrita, oral o de evaluación de procesos o productos a través de la

observación. Ellas permiten disminuir el tiempo y los problemas que se les plantean a los docentes en el momento de la corrección, les dan mayor seguridad y contribuyen a mejorar la confiabilidad de los resultados de la evaluación. Se deben emplear con precaución para no destruir el espíritu de estas pruebas que es la libertad de presentación y organización que se le brinda al alumno. De otro modo sería preferible utilizar instrumentos de tipo objetivo, esto es, instrumentos que admiten una y sólo una respuesta correcta y no diferentes calidades de respuesta que son más difíciles de analizar.

Economía

Es importante, como hemos dicho antes, que el programa de evaluación sea económico en tiempo, trabajo y costo. Es conveniente examinar previamente estas características para determinar la conveniencia de su empleo.

INTEGRACIÓN DE LOS TRES ASPECTOS

Los tres aspectos relativos a la practicidad del programa o de cada instrumento deben considerarse conjuntamente para lograr una evaluación integral de este rasgo.

Cada aspecto presenta ventajas y desventajas que se compensan o equilibran cuando se analizan los otros dos. Un costo relativamente elevado, por ejemplo, podría ser compensado por un alto grado de administrabilidad, o un tiempo prolongado dedicado a la construcción de un instrumento puede ser equilibrado por un tiempo menor destinado a la corrección e interpretación de los resultados. Del mismo modo, un tiempo extenso de construcción

puede compensarse por la facilidad de administración simultánea a un grupo muy numeroso de alumnos.

La practicidad no es, seguramente, la característica más importante de un programa o de un instrumento de evaluación. Pero sí adquiere relevancia cuando de ella depende la frecuencia de la evaluación. Se observa que, por lo general, cuando el programa es práctico o cuando el instrumento es de fácil administración e interpretación se evalúa más frecuentemente y ello contribuye a estructurar mejor los procesos de enseñanza y los procesos de aprendizaje de los alumnos por parte del docente y de los mismos alumnos.

UTILIDAD

Íntimamente asociada a las otras tres características que mencionamos - validez, confiabilidad y practicidad-, la utilidad de una prueba resulta de su capacidad para satisfacer las necesidades específicas relacionadas con los procesos de enseñanza y aprendizaje. Si los resultados de la evaluación no son útiles para la orientación de los alumnos, del docente, de la escuela como institución, de los padres de los alumnos, de los sectores interesados en la calidad de la educación, poco importa que, en un marco restringido, el programa y cada instrumento sean confiables y prácticos. En lo que respecta a la validez, ya hemos dicho que es relativa a los criterios y, tratándose de públicos diferentes, dichos criterios pueden ser heterogéneos. Para que la evaluación sea útil, la validez para alguno o algunos de esos públicos es indispensable. La pregunta esencial en el diseño de un programa de evaluación es, por lo tanto, ¿qué usos se les dará a los resultados obtenidos? Son los implicados en los procesos de evaluación quienes deben dar las respuesta

Editorial

La investigación sobre la planeación educativa □

El inicio del año representa la oportunidad de renovar con nuevos bríos el ciclo de actividades cotidianas que nos ocuparon hasta el presente y que seguirán en el centro de nuestras inquietudes. El proyecto editorial de Perfiles Educativos se congratula en retomar sus tareas correspondientes a este 2011, e invita a todos sus autores, dictaminadores y lectores, así como a su cuerpo directivo, administrativo y técnico, a continuar con la producción y difusión entusiasta de trabajos de investigación educativa para contribuir al conocimiento sistemático y al desarrollo de la educación formal y de las disciplinas que le dan sustento.

En esta ocasión compartimos una breve reflexión sobre la planeación educativa, elemento que puede jugar un papel importante en el funcionamiento general del sistema educativo, tanto si se piensa en la programación macro como en el nivel micro, como auxiliar en el desempeño de la labor docente en los escenarios previstos por la organización escolar. En ambos casos se trata de elementos clave a investigar, sobre todo en la perspectiva de comprender mejor desde el diseño hasta la puesta en marcha de la planeación, y de orientar las acciones futuras.

En sentido general, la planeación dirigida al sistema educativo en su conjunto se expresa en un cuerpo de políticas que intentan traducirse en orientaciones para la acción de los agentes directamente involucrados en el funcionamiento del sistema escolar; por lo común cada nueva administración define las políticas que se identificarán con su proyecto de desarrollo y es frecuente que no se considere la situación actual, fruto de las políticas anteriores, para poner en marcha las nuevas.

La investigación educativa siempre ha estado presente en el intento de valorar el impacto de las políticas, sin embargo, debido en parte por la magnitud y complejidad de la tarea, no siempre se ha contado con el panorama global de los efectos esperados e inesperados de las mismas, es decir, con un balance equilibrado que arroje luz para las acciones futuras y consolide el conocimiento social disponible sobre lo experimentado. En este sentido, convendría alentar a las nuevas generaciones de

La información general de este editorial fue tomada de: M.C. Wittrock, 1990.

Perfiles Educativos | vol. XXXIII, núm. 131, 2011 | IISUE-UNAM

Mario Rueda Beltrán | Editorial

investigadores a que orientaran sus esfuerzos a la producción de conocimiento sistemático acerca de los distintos planteamientos generados por los diferentes cuerpos directivos, con la idea de contar con información documentada acerca del ciclo que comprende la planeación, la puesta en marcha, la evaluación y finalmente una planeación renovada. Iniciativas de este tipo también podrían tener repercusiones en los encargados de definir las políticas, al suministrarles datos para la valoración de los impactos asociados a la implementación de lo planificado con el propósito de dar continuidad a los aspectos positivos y evitar la continuidad o repetición de lo revelado como ineficaz o impertinente.

La planeación es un elemento indispensable para la orientación de todas las acciones vinculadas con la organización escolar, más aún ahora en que las exigencias derivadas de grandes cambios sociales se vuelcan hacia la escuela en busca de la definición y el cumplimiento de nuevas funciones, así como al señalamiento de graves problemas ya identificados con anterioridad, pero ahora urgidos con mayor fuerza de una pronta solución, como el

incumplimiento de una cobertura universal, los bajos niveles de egreso, el abandono escolar, la desigualdad de oportunidades escolares, la falta de pertinencia de los aprendizajes para desenvolverse exitosamente en las sociedades actuales y la amplitud creciente de la brecha digital, entre otros.

En este contexto, la planeación desarrollada por el maestro puede representar la diferencia en el cumplimiento de las metas socialmente esperadas en la formación de los futuros profesionales y ciudadanos. Una planeación que vaya más allá de las exigencias administrativas y que medie entre los requerimientos diseñados por directivos y especialistas del currículo, y las características de cada grupo particular en convivencia con un maestro con características personales y profesionales propias y una materia a enseñar también con requerimientos específicos para su apropiación plena.

La investigación educativa, sobre todo la que se desarrolla con los maestros del nivel primario, arroja información que deja en claro la existencia de distintos tipos y funciones de la planeación (Wittrock, 1990: 454-474). Adicionalmente se ha identificado que los tipos de planeación también varían si se trata de profesores principiantes o experimentados. Los tipos de planeación observados aluden a distintos lapsos de tiempo considerados: anual, semanal, diaria, entre otros, o bien a los contenidos, como la planificación de una unidad de aprendizaje o lección. En diversos estudios realizados con maestros del nivel básico se observó que invierten sus esfuerzos de planeación para la estructuración, organización y administración de periodos limitados de instrucción en la sala de clase. Asimismo, se reporta como uno de los productos de la planeación las rutinas, lo que pone en evidencia que esta actividad obedece, al

Mario Rueda Beltrán | Editorial

Menos en parte, a las presiones a favor de la simplificación y la administración eficaz del tiempo, además de que sirve para coordinar las secuencias específicas de conducta tanto del propio docente como de los alumnos. El tipo de planeación varía también a lo largo del año escolar y puede dar una mayor importancia relativa a la preparación del ambiente físico del aula, la valoración de las capacidades de los alumnos o al establecimiento del sistema social del salón de clase.

Las razones aludidas por un conjunto de maestros sobre el porqué de la planificación se refieren a la satisfacción de necesidades personales inmediatas, como la disminución de la incertidumbre o el aumento de la confianza; a la posibilidad de acceder a la instrucción, como aprenderse el material u organizar el flujo de actividad del grupo; o bien, al señalamiento de una función directa, como poner en marcha una actividad u ofrecer un marco a la instrucción y la evaluación. Algunos de los maestros estudiados utilizan la planificación para cumplir con el requisito administrativo de presentar sus planes al director de la escuela, y también para los casos en los que se requieren planes especiales en previsión de la participación de maestros sustitutos.

En el conjunto de estudios realizados se muestra también que la planeación les sirve a los maestros para adaptar los contenidos y actividades previstos en el currículo al grupo particular de alumnos que tienen a su cargo, y que las revisiones anuales de lo ocurrido los ubican en una perspectiva de mayor certeza sobre los contenidos a cubrir y sobre cómo hacerlo, ya que toman como punto de partida las secuencias y ritmos de enseñanza de la experiencia del ciclo escolar anterior para hacer los ajustes que se requieren.

Las evidencias de los estudios estadounidenses dejan constancia de algunas de las distintas modalidades que adoptan los maestros en su interacción con los planes de estudio oficiales y las acciones de planeación que, como ya se dijo, pueden servir para revisar a mediano plazo el currículo, conocer las distintas formas de adaptación del mismo en contextos muy variables, consolidar la seguridad del docente al proporcionarle una guía para su conducción del proceso de enseñanza y contar con elementos para la mejora continua, entre otras. ¿Cuáles son los tipos de planeación de los maestros en contextos latinoamericanos?, ¿cuáles son las funciones que cubren estas acciones de planeación? Los investigadores educativos y los estudiantes en formación en posgrados en educación tienen un tema sugerente a considerar en sus próximas agendas de trabajo.

Otros esfuerzos derivados también de los estudios sobre la planificación del docente han estado encaminados a la creación de modelos descriptivos del proceso de planificación; algunas de las investigaciones han arrojado información distinta de aquella que se deriva de uno de los

Perfiles Educativos | vol. XXXIII, núm. 131, 2011 | IISUE-UNAM

Mario Rueda Beltrán | Editorial

Modelos más influyentes y que postula para la planeación la secuencia: definición de objetivos, selección de actividades de aprendizaje, organización de dichas actividades y especificación de los procedimientos de evaluación. Los testimonios obtenidos muestran que estos elementos no son considerados por todos los maestros ni en el orden propuesto ni en la consideración de su importancia relativa; nuevamente la diferencia se hace

presente entre los profesores principiantes y los experimentados. Inclusive se han detectado modelos en los que la planeación se concentra, en orden de importancia decreciente, en las necesidades, capacidades e intereses de los alumnos, la asignatura, las metas y los métodos de enseñanza; y en los que se otorga poca importancia a la evaluación, además de que denotan un escaso interés en la vinculación entre el curso planificado y el plan de estudios general.

La información anterior sólo tiene la pretensión de llamar la atención sobre un tema relevante de investigación que puede contribuir al seguimiento de propuestas de innovación educativa en las cuales, sin duda, se pondrá en juego el concepto mismo de la planeación y las funciones que se le atribuyen en cada contexto. La investigación educativa sobre la planeación ha empleado hasta ahora, entre otros recursos metodológicos para su estudio, la observación, la entrevista, el diario, el pensamiento en voz alta, la estimulación del recuerdo y la etnografía, que constituyen una plataforma suficiente para alentar el desarrollo de trabajos latinoamericanos sobre el tema. Los resultados que se obtengan de este tipo de investigaciones, además de completar el conocimiento universal disponible sobre la planeación, podrán coadyuvar al conocimiento sistemático y a la comprensión de una actividad que puede reducir la incertidumbre acerca de la interacción de docentes y alumnos, así como organizar y orientar las acciones del proceso de enseñanza y aprendizaje y proporcionar la oportunidad de su revisión constante. Todo ello ante el reconocimiento de que se trata de un proceso social complejo, la docencia, que se caracteriza por imprevistos, interrupciones y digresiones, resultantes de la interacción de grupos de estudiantes heterogéneos, maestros con estilos de enseñanza muy diversos y condiciones institucionales disímiles. La investigación educativa puede contribuir con su conocimiento para que la labor del profesorado sea más

efectiva y pertinente, sobre todo en una sociedad que ve en la educación formal una estrategia para sembrar un mejor futuro.

Mario Rueda Beltrán

REFERENCIA

WITTROCK, M.C. (1990), La investigación de la enseñanza III. Profesores y alumnos, Barcelona, Paidós Educador.

Perfiles Educativos | vol. XXXIII, núm. 131, 2011 | IISUE-UNAM

Mario Rueda Beltrán | Editorial

I.3 PLANEACIÓN DEL PROCESO DE EVALUACIÓN DE INSTITUCIONES EDUCATIVA.

La planeación educativa: ¿ciencia o política?*

A pesar de que fue recibida con entusiasmo y se difundió por todas partes, la planeación educativa no ha tenido mucho éxito en América Latina. El término “éxito” requiere alguna definición. El sentido más exigente sería el de que, mediante la implantación fiel de un plan (o proceso de planeación) educativo, se ayudara efectivamente a resolver los grandes problemas y eliminar las injusticias que padecen los pueblos latinoamericanos. Si adoptáramos como criterio para

juzgar la eficiencia de la planeación educativa cambios radicales en las estructuras económicas y políticas de las naciones de América Latina, tendríamos dificultad en poder señalar casos de éxito. Hasta el momento, en casi todas partes la educación ha sido vista como reproductora de estructuras existentes, y la planeación educativa ha reflejado la hegemonía de gobiernos interesados en mantener el statu quo. Vista desde esta óptica, la planeación educativa no puede servir mayormente hasta que haya otros cambios, y en cual caso su análisis no sería de mayor interés.

Este trabajo no evalúa entonces la capacidad de la planeación educativa para producir por sí misma cambios radicales en la sociedad. Al contrario, se enfoca en la utilización de la planeación educativa para reforzar estructuras y procesos existentes, o para facilitar procesos que favorecen el mantenimiento de estructuras amenazadas por cambios en sus ambientes externos o debido a procesos internos. En este caso, la definición de “éxito” del proceso de planeación está en su contribución al mantenimiento de estructuras vigentes.

Aun con esta definición, la evaluación del “éxito” de la planeación es difícil, pues solamente con el tiempo sería posible ver los efectos de la implantación de cambios inducidos o impuestos por un proceso de planeación. Por tanto, hemos adoptado para este trabajo un criterio más flexible pero aun estricto con respecto a la mayoría de las experiencias en planeación educativa en América

Latina. Tomamos como criterio de “éxito” de un proceso de planeación, la implantación de los planes educativos en el sistema educativo.

Este criterio es aún estricto porque constantemente se ha hablado mucho sobre la reforma educativa en casi todos los países, pero pocas han llegado a ser implantadas; incluso en Cuba, el caso más radical de reforma, la planeación revistió inicialmente escasa importancia. En la mayoría de los países es posible que los planes educativos llenen los anaqueles, pero una vez publicados generalmente sólo los consultan los investigadores académicos.²

Con frecuencia, las oficinas de planeación educativa no hacen más que recopilar y publicar estadísticas o patrocinar investigaciones sobre interesantes proyectos piloto que nunca se hacen extensivos a todo el sistema. Raras veces la Oficina de Planeación tiene control del presupuesto o de la asignación de los recursos, la prueba más evidente de su fuerza y su importancia.³

El estudio de caso que presentamos en este trabajo explora el desarrollo de una experiencia “exitosa” de planeación educativa.

En este caso, una oficina de planeación educativa no sólo produjo planes que consultaban regularmente, sino que también tuvo un estrecho control de la formulación del presupuesto y de las políticas educativas generales.

Dada esta definición del problema a estudiar, nuestra atención a la planeación se enfoca en el poder. El “éxito” de la planeación depende de la capacidad de un grupo u organización de imponer sus ideas sobre otros grupos y

organizaciones. La calidad de las ideas puede influir en la generación y ejercicio del poder, pero no serán esas ideas en sí la explicación del “éxito” en la implantación del plan, sino que dependerá de otros factores.

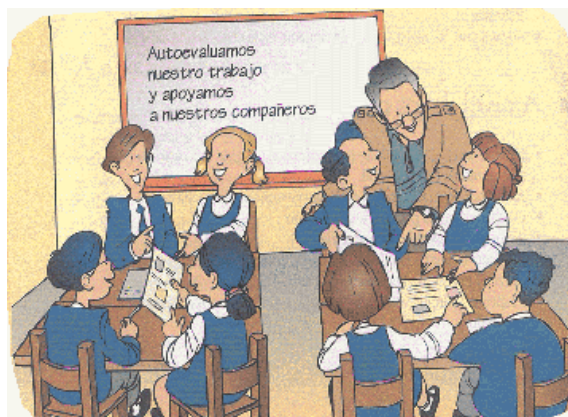
Consideramos que el éxito de la planeación, definido de esta manera, se da en función de las habilidades en el desarrollo y manejo de las relaciones políticas, y no de la competencia técnica. Consideramos la planeación esencialmente como proceso político, en que se manipula información empleando técnicas con una alta carga de legitimidad científica, para imponer nuevas formas de organización. El auge de la planeación educativa en América Latina se explica, creemos, en términos del interés de nuevos grupos tecnócratas de competir por el poder con grupos dominantes tradicionales. A la vez, la planeación ha creado grandes expectativas de cambios y, de esta manera, ha servido para mantener por un tiempo más la estructura básica del sistema político y económico.

I.I INSTRUMENTOS PARA LA EVALUACIÓN DE UNA INSTITUCIÓN EDUCATIVA.

Por Noemí Álvarez Vallina

Se evalúa siempre para tomar decisiones. No basta con recoger información sobre los resultados del proceso educativo y emitir únicamente un tipo de calificación, si no se toma alguna decisión, no existe una auténtica evaluación.

Así pues, la evaluación es una actividad o proceso sistemático de identificación, recogida o tratamiento de datos sobre elementos o hechos educativos, con el objetivo de valorarlos primero y, sobre dicha valoración, tomar decisiones (García Ramos, 1989).



La evaluación, por tanto, se caracteriza como: Un proceso que implica recogida de información con una posterior interpretación en función del contraste con determinadas instancias de referencia o patrones de deseabilidad, para hacer posible la emisión de un juicio de valor que permita orientar la acción o la toma de decisiones.

Finalmente, deben diferenciarse los conceptos de investigación y evaluación. Ambos procesos tienen muchos elementos comunes, aunque se diferencian en sus fines:

- La evaluación es un proceso que busca información para la valoración y la toma de decisiones inmediata. Se centra en un fenómeno particular. No pretende generalizar a otras situaciones.
- La investigación es un procedimiento que busca conocimiento generalizable, conclusiones (principios, leyes y teorías), no tiene necesariamente una aplicación inmediata (De la Orden, 1989).

ÁMBITOS DE LA EVALUACIÓN

Tradicionalmente, la evaluación se ha venido aplicando casi con exclusividad al rendimiento de los alumnos, a los contenidos referidos a conceptos, hechos, principios, etc., adquiridos por ellos en los procesos de enseñanza. A partir de los años sesenta, la evaluación se ha extendido a otros ámbitos educativos: actitudes, destrezas, programas educativos, materiales curriculares didácticos, la práctica docente, los centros escolares, el sistema educativo en su conjunto y la propia evaluación.

<u>PARA EL ALUMNO</u>	<ol style="list-style-type: none"> 1. Informativa. 2. Orientadora. 3. Motivadora.
<u>PARA EL PROFESOR</u>	<ol style="list-style-type: none"> 1. Conocer el ritmo de aprendizaje. 2. Diagnosticar su nivel. 3. Pronosticar. 4. Agrupar o clasificar. 5. Motivar a los alumnos 6. Calificar. 7. Mantener informados a los padres.
<u>COMO RECURSO DE ACCIÓN TUTORIAL</u>	<ol style="list-style-type: none"> 1. Finalidad de orientación escolar, profesional y personal del alumno.
<u>PARA EL PROCESO EDUCATIVO</u>	<ol style="list-style-type: none"> 1. Constatación y revisión de forma continúa de la validez en la actuación de cada uno de los diferentes elementos que intervienen incluso en la evaluación. (Metaevaluación).

Esta extensión de la evaluación a otros ámbitos tuvo lugar en los Estados Unidos a finales de los años 50 debido a circunstancias tales como: la crítica a la eficacia de las escuelas públicas, la gran inversión dedicada a la educación que exigía una rendición de cuentas (Accountability), etc. Por tanto, el campo de aplicación de la evaluación se extiende a alumnos, profesores, directivos, instituciones, la administración, etc. Y va a ser, precisamente, a raíz de la extensión del ámbito evaluador cuando van a surgir una serie de modelos de evaluación de gran relevancia.

EVALUACIÓN/PROMOCIÓN

La decisión de promoción es la que, con más frecuencia, debe enfrentar el profesor, desde las promociones formales (curso a curso) hasta las promociones diarias (de una tarea a otra, cuando se considera que se ha alcanzado un nivel de conocimientos suficiente).

Por tanto, la evaluación puede resultar un elemento estimulante para la educación en la medida en que pueda desembocar en decisiones de promoción positivas, y para ello es preciso que el sistema educativo sea público y coherente, ofreciendo la información precisa para ofrecen la dificultades que puedan surgir. Para ello, es necesaria una definición clara de los objetivos previos y una recuperación inmediata en caso de fracaso. En caso de que el fracaso sea reiterado, se hace imprescindible la utilización de procesos diagnósticos y terapéuticos. Por tanto, lo deseable es la promoción tanto desde el punto de vista del aprendizaje como desde el punto de vista del desarrollo armónico de la persona.

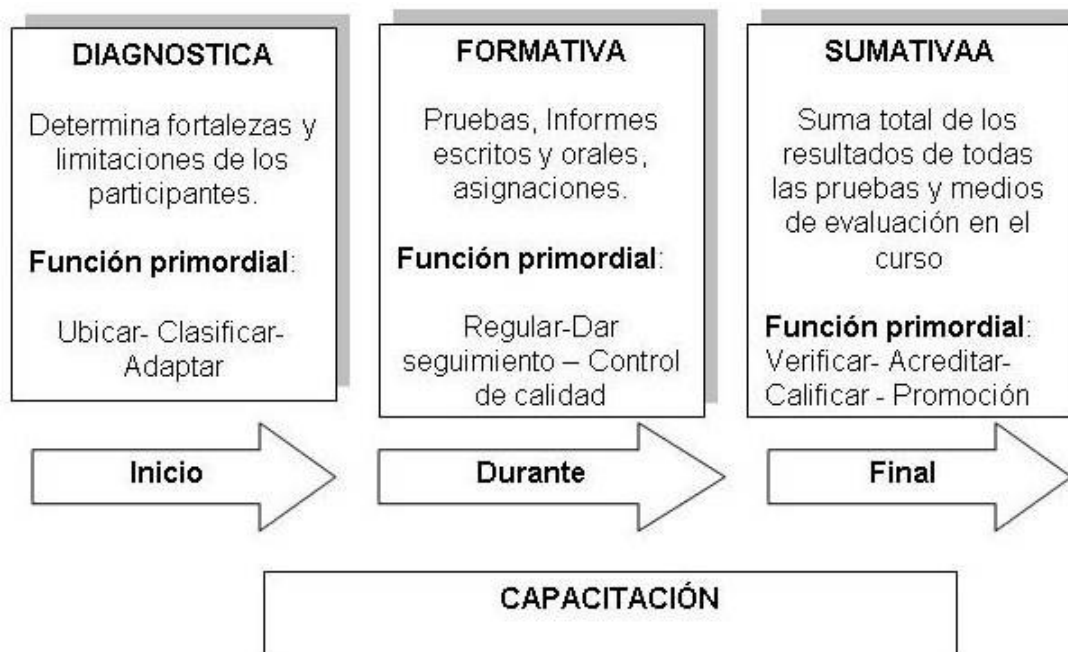
TIPOS DE EVALUACIÓN

Esta clasificación atiende a diferentes criterios. Por tanto, se emplean uno u otro en función del propósito de la evaluación, a los impulsores o ejecutores de la misma, a cada situación concreta, a los recursos con los que contemos, a los destinatarios del informe evaluador y a otros factores.

Según su finalidad y función

- a) **Función formativa:** la evaluación se utiliza preferentemente como estrategia de mejora y para ajustar sobre la marcha, los procesos educativos de cara a conseguir las metas u objetivos previstos. Es la más apropiada para la evaluación de procesos, aunque también es formativa la evaluación de productos educativos, siempre que sus resultados se empleen para la mejor de los mismos. Suele identificarse con la evaluación continua.

- b) **Función sumativa:** suele aplicarse más en la evaluación de productos, es decir, de procesos terminados, con realizaciones precisas y valorables. Con la evaluación no se pretende modificar, ajustar o mejorar el objeto de la evaluación, sino simplemente determinar su valía, en función del empleo que se desea hacer del mismo posteriormente.



Según su extensión

- a) **Evaluación global:** se pretende abarcar todos los componentes o dimensiones del alumno, del centro educativo, del programa, etc. Se considera el objeto de la evaluación de un modo holístico, como una totalidad interactuante, en la que cualquier modificación en uno de sus componentes o dimensiones tiene consecuencias en el resto. Con este tipo de evaluación, la comprensión de la realidad evaluada aumenta, pero no siempre es necesaria o posible. El modelo más conocido es el CIPP de Stufflebeam.
- b) **Evaluación parcial:** pretende el estudio o valoración de determinados componentes o dimensiones de un centro, de un programa educativo, de rendimiento de un alumno, etc

2. Tipos de evaluación

FUNCION	FORMATIVA		SUMATIVA
TIEMPO	INICIAL	CONTINUA	FINAL
REFERENTE	NORMATIVA	CRITERIAL	PESONALI- ZADA
AGENTE	AUTOEVA- LUACION		HETERO- EVALUACION

Según los agentes evaluadores

a) Evaluación interna: es aquella que es llevada a cabo y promovida por los propios integrantes de un centro, un programa educativo, etc.

A su vez, la evaluación interna ofrece diversas alternativas de realización: autoevaluación, heteroevaluación y coevaluación.

- Autoevaluación: los evaluadores evalúan su propio trabajo (un alumno su rendimiento, un centro o programa su propio funcionamiento, etc). Los roles de evaluador y evaluado coinciden en las mismas personas.
- Heteroevaluación: evalúan una actividad, objeto o producto, evaluadores distintos a las personas evaluadas (el Consejo Escolar al Claustro de profesores, un profesor a sus alumnos, etc.)

- **Evaluación:** es aquella en la que unos sujetos o grupos se evalúan mutuamente (alumnos y profesores mutuamente, unos y otros equipos docentes, el equipo directivo al Consejo Escolar y viceversa). Evaluadores y evaluados intercambian su papel alternativamente.

b) **Evaluación externa:** se da cuando agentes no integrantes de un centro escolar o de un programa evalúan su funcionamiento. Suele ser el caso de la "evaluación de expertos". Estos evaluadores pueden ser inspectores de evaluación, miembros de la Administración, investigadores, equipos de apoyo a la escuela, etc.

	INICIAL/ DIAGNOSTICO	REFERENCIA PARA EL PROCESO E-A DEL ALUMNO
FUNCION	FORMATIVA	RETROALIMENTACION DURANTE EL PROCESO EDUCATIVO
	SUMATIVA	DETERMINACION DEL PRODUCTO FINAL EN REFERENCIA A LA EVALUACION INICIAL
REFERENTE	NORMATIVA	EVALUACION DEL ALUMNO EN FUNCION DEL RENDIMIENTO DEL GRUPO DE CLASE
	CRITERIAL	EVALUACION DEL ALUMNO A PARTIR DE UNOS CRITERIOS EXTERNOS
MOMENTO	COMIENZO	AL INICIO DEL PROCESO DE E-A, PARA CONOCER LOS CONOCIMIENTOS PREVIOS
	PROCESO	DURANTE EL PROCESO DE E-A PARA REAJUSTAR, MEJORAR Y PERFECCIONAR
	FINAL	AL TERMINO DEL PROCESO DE E-A PARA CONOCER LOS CONOCIMIENTOS ALCANZADOS
AGENTE	AUTOEVALUACION	EL ALUMNO EVALUA SUS PROPIAS ACTUACIONES
	COEVALUACION	LOS ALUMNOS SE EVALUAN POR PAREJAS O ENTRE GRUPOS
	HETEROEVALUACION	EL ALUMNO ES EVALUADO POR EL PROFESOR Y POR LAS PERSONAS RELACIONADAS

Estos dos tipos de evaluación son muy necesarios y se complementan mutuamente. En el caso de la evaluación de centro, sobre todo, se están extendiendo la figura del "asesor externo", que permite que el propio centro o programa se evalúe a sí mismo, pero le ofrece su asesoría técnica y cierta objetividad por su no implicación en la vida del centro.

Según el momento de aplicación

a) Evaluación inicial: se realiza al comienzo del curso académico, de la implantación de un programa educativo, del funcionamiento de una institución escolar, etc. Consiste en la recogida de datos en la situación de partida. Es imprescindible para iniciar cualquier cambio educativo, para decidir los objetivos que se pueden y deben conseguir y también para valorar si al final de un proceso, los resultados son satisfactorios o insatisfactorios.

b) Evaluación procesual: consiste en la valoración a través de la recogida continua y sistemática de datos, del funcionamiento de un centro, de un programa educativo, del proceso de aprendizaje de un alumno, de la eficacia de un profesor, etc. a lo largo del periodo de tiempo fijado para la consecución de unas metas u objetivos. La evaluación procesual es de gran importancia dentro de una concepción formativa de la evaluación, porque permite tomar decisiones de mejora sobre la marcha.

c) Evaluación final: consiste en la recogida y valoración de unos datos al finalizar un periodo de tiempo previsto para la realización de un aprendizaje, un programa, un trabajo, un curso escolar, etc. o para la consecución de unos objetivos.

Según el criterio de comparación

Cualquier valoración se hace siempre comparando el objeto de evaluación con un patrón o criterio. En este sentido, se pueden distinguir dos situaciones distintas:

- a) En caso de que la referencia sea el propio sujeto (sus capacidades e intereses, las metas que se había propuesto alcanzar, considerando el tiempo y el esfuerzo invertidos por el sujeto, y teniendo en cuenta sus aprendizajes previos) o cualquier otro objeto de la evaluación en si mismo (las características de partida de un programa, los logros educativos de un centro en el pasado, etc.), estaremos empleando la AUTOREFERENCIA como sistema
- b) En el caso de que las referencias no sean el propio sujeto, centro, programa, etc., lo que se conoce como HETEROREFERENCIA, nos encontramos con dos posibilidades:

b.1) Referencia o evaluación criterial:

Aquella en las que se comparan los resultados de un proceso educativo cualquiera con los objetivos previamente fijados, o bien con unos patrones de realización, con un conjunto de situaciones deseables y previamente establecidas. Es el caso en el que comparamos el rendimiento del alumno con los objetivos que debería haber alcanzado en un determinado plazo de tiempo, o los resultados de un programa de educación compensatoria con los objetivos que éste se había marcado, y no con los resultados de otro programa.

b.2) Referencia o evaluación normativa:

El referente de comparación es el nivel general de un grupo normativo determinado (otros alumnos, centros, programas o profesores).

Lo correcto es conjugar siempre ambos criterio para realizar una valoración adecuada, aunque en el caso de la evaluación de alumnos, nos parece siempre más apropiada la evaluación que emplea la autorreferencia o la evaluación criterial. El empleo de uno u otro tipo de evaluación dependerá siempre de los propósitos de la evaluación y de su adecuación al objeto de nuestra evaluación.

NOEMI ALVAREZ VALLINA