



# *La evaluación cualitativa de planes y centros de perfeccionamiento de profesorado: Una forma de mejorar la profesionalidad docente*

Miguel A. Santos Guerra (\*)  
*Dpto. Didáctica y Organización Escolar*  
*Universidad de Málaga*

## RESUMEN

*El autor plantea la necesidad de realizar la evaluación de los procesos educativos institucionales de perfeccionamiento del profesorado no como un apéndice final de la actividad, sino como un análisis valorativo riguroso en el curso vivo de la actividad, facilitando la toma de decisiones y produciendo una mayor comprensión de los fenómenos.*

*Focaliza la atención en la evaluación cualitativa como una exigencia de la naturaleza compleja, dinámica, absolutamente peculiar (por lo tanto imprescindible) y cargada de valores del fenómeno educativo que tiene lugar en los programas y en los Centros de perfeccionamiento de profesores.*

*Se atiende a los planteamientos metodológicos coherentes con la filosofía del enfoque y se establecen las condiciones técnicas, sociales, políticas y éticas que se derivan de la entraña de la concepción naturalista de la evaluación.*

*Finalmente, el autor se detiene a considerar las fuentes de la credibilidad que han de tenerse en cuenta al plantear el proceso de metaevaluación.*

La evaluación cualitativa es un proceso que permite poner en marcha estrategias rigurosas de reflexión y de valoración sobre ac-

tividades en desarrollo. Descartamos en nuestro enfoque la concepción de un tipo de evaluación que se sitúa como un apéndice fi-

(\*) Plaza del Ejido, s/n.  
29013 Málaga

nal del programa con la pretensión de medir los resultados obtenidos por el mismo, ya que este planteamiento nos impide analizar la evolución del programa, el funcionamiento de la caja negra donde se desarrollan los procesos, la dinámica interna que se establece entre los protagonistas de la actividad y las intenciones, los motivos y las valoraciones de quienes están llevando a cabo el plan.

Se evalúa para comprender la naturaleza de los programas, para mejorar a través del análisis y de los resultados lo que se hace y lo que se pretende hacer y para generar y reformular teorías interpretativas de la realidad. La evaluación que proponemos no tiene solamente la pretensión de etiquetar un programa como bueno/malo, ni siquiera la de facilitar la toma racional de las decisiones sobre otros programas o la continuidad y mejora del que se ha evaluado sino que se convierte en un instrumento de la formulación teórica, en un proceso dialéctico permanente entre la racionalidad teórica y la racionalidad práctica.

Existen, en esencia, dos modelos de evaluación de programas instituciones; uno, que llamaremos experimental, que se basa en el control inicial, la aplicación del plan y el análisis final que permita *comprobar* qué modificaciones ha producido la actividad en cuestión. Los instrumentos que utiliza son de carácter cuantitativo y pretende establecer un control de todas las variables que intervienen, con el fin de establecer nexos causales (o meramente correlacionales) entre el programa y los efectos pretendidos del mismo en el momento en que finaliza. Los sujetos que protagonizan la actividad no intervienen en la evaluación más que para comprobar los efectos que ha tenido en ellos el programa a juicio del evaluador externo, que opera como una instancia sancionadora.

El logro de los objetivos propuestos será valorado mediante la comprobación del grado de consecución de los mismos alcanzado

al finalizar el programa, sin cuestionarse la bondad y adecuación de los objetivos propuestos, los procesos mediante los cuales se ha intentado conseguirlos, la correspondencia medios/fines, los efectos secundarios y no pretendidos, los efectos sutiles, la valoración de los mismos sujetos, etc.

«Los beneficios pueden resultar difusos, demorarse durante largo tiempo o incluso quedar más allá del escrutinio de los evaluadores» (Stake, 1980).

La evaluación experimental descansa en una concepción positivista de la realidad y de la ciencia y se afana en defender una objetividad y un rigor en las mediciones, alejándose lo más posible de la subjetividad del evaluador y de los participantes en el programa. Este planteamiento encierra en su entraña una discutible concepción de la actividad humana (y de la actividad educativa, en concreto) que se equipara al comportamiento de los fenómenos físicos de la naturaleza. Sin embargo, el comportamiento educativo está transido de ambigüedad, de valores, de imprevisibilidad, de motivos, intenciones, expectativas, etc. difícilmente domeñables a los planteamientos del enfoque positivista.

«Si se consideran de esta manera las acciones humanas, es claro que cualquier intento de explicarlas del mismo modo las ciencias naturales explican el comportamiento de los objetivos naturales, priva a aquellas de sus significados propios, que reemplaza por las interpretaciones causales del tipo que demanda el concepto positivista de explicación. Cuando esto ocurre las acciones significativas se reducen a patrones de conducta que, como la dilatación de los metales, se suponen determinados por fuerzas externas y pueden reducirse a la explicación científica convencional. La acción queda desprovista de su sentido y halla su lugar en un cálculo de movimientos que sólo tienen el sentido ilícito que les dan los significados y valoraciones que el científico positivista trata en vano de expulsar de sus teorías. Si se quiere evitar esto, si el intento de comprender los fenómenos humanos y sociales han de tomarse en serio, es preciso admitir que las ciencias sociales versan sobre una materia temática totalmente diferente de la de las ciencias naturales, y que los métodos y la forma de explicación que se utilicen en ambos tipos de ciencia han de ser completamente distintos» (Carr y Kemmis, 1988)

La naturaleza de los fenómenos que se evalúan exige un enfoque adecuado. Si tuviera que estudiar el movimiento de las bolas de billar sobre la mesa de juego, debería medir la longitud de los desplazamientos, la fuerza de los golpes, la resistencia de la superficie, etc. Para ello será necesario utilizar medidas precisas y fiables. Pero si quisiera conocer por qué juegan al billar las personas que rodean la mesa, no conocería la verdadera significación de las intenciones y del comportamiento si me limitase a aplicar un instrumento que *midiera* los motivos por los que están jugando. El problema no estriba solamente en que esa medición no pueda ser realizada con rigor sino en que su mismo intento desvirtúa la naturaleza del fenómeno y nos aleja de la comprensión de su verdadero significado.

Valorar el éxito de un programa, el aprovechamiento de un curso, el funcionamiento de un Centro de Profesores exige un punto de partida que respete la naturaleza del fenómeno educativo y que tenga en cuenta su extraordinaria complejidad, dado el polivalente entramado de variables individuales, políticas, económicas y sociales que lo integran.

El enemigo más importante de la complejidad de una tarea es la simplificación de la misma, la reducción de su dificultad al esquematismo y a la trivialización y, desde luego, el simplismo de la apariencia de rigor que esconde el verdadero significado de la realidad. *La magia de los números* (Siedman, 1977) encierra el peligro del embaucamiento científico. La esquematización que supone una definición operativa de objetivos (Gimeno Sacristán, 1984), la utilización de unos instrumentos de medida estandarizados y la aplicación de paquetes estadísticos que muestran la significatividad estocástica de los resultados, impide una argumentación, un análisis riguroso, una crítica elaborada, una reflexión profunda, una interpretación cualitativa de la realidad.

«Nuestra búsqueda de la racionalidad científica y la evidencia factual nos ha llevado a una reverencia irracional por la autoridad de los números, raramente cuestionada» (Siedman, 1977)..

Limitarse a la comprobación del grado de logro de unos objetivos previos puede hacer olvidar la argumentación profunda sobre el significado y la importancia de los mismos, sobre su adecuación a las necesidades de los individuos, sobre los procesos desarrollados para conseguirlos y sobre los efectos no pretendidos en el diseño del programa. La simplificación de los resultados y del proceso de evaluación resulta preocupante no sólo por la parcialidad y escasez de lo que muestra sino por la importancia de lo que oculta.

«El valor de un programa depende de su relevancia para otros cursos o programas, de sus efectos en la moral del profesor, de su compatibilidad con la economía de la comunidad, etc., etc.» (Stake, 1978).

El uso de diseños experimentales para efectuar la evaluación de un programa educativo de formación de profesorado parte de un diseño que el evaluador sigue con rigor (cortando el camino de lo imprevisible), de unas hipótesis previas que condicionan la exploración, no tiene en cuenta el contexto, parte del presupuesto de que los programas educativos son repetibles, cuantifica los resultados, establece comparaciones a través de la significatividad estadística, atribuye nexos causales entre variables controladas en situaciones de *laboratorio* y plantea generalizaciones para toda la población desde el estudio de muestras representativas. Este modelo está presidido por la *objetividad* del diseño y de la medida, por la eliminación de la visión subjetiva de los sujetos y de los propios evaluadores, por una pretendida neutralidad ideológica, política e incluso teórica...

Nosotros abogamos por un modelo de evaluación de carácter diferente, asentado sobre otros presupuestos teóricos y metodológicos:

### **Evaluación formativa, con repercusión en la realidad de los programas**

Plantear la evaluación como un elemento externo, añadido y terminal hace que pierda una de sus mayores virtualidades, a saber, la repercusión sobre el propio programa o sobre el funcionamiento del Centro evaluado. Los participantes reflexionan sobre la actividad, los evaluadores se convierten en instrumentos encauzadores y dinamizadores de la crítica y de la argumentación interactiva, de tal modo que los integrantes del programa se cuestionan, en la dialéctica permanente de los datos y las interpretaciones, de la teoría y la práctica, la naturaleza de aquello que realizan y las repercusiones que desencadena el actuar de una forma determinada en ese contexto preciso.

«La evaluación formativa es el intento de utilizar el reaprovechamiento para formar programas progresivos y para ayudar a los participantes a conseguir sus objetivos. Así, hay una relación recíproca en la evaluación formativa, entre los evaluadores externos y el personal responsable de la consecución del programa. Los evaluadores tienen la responsabilidad de ayudar a que el programa tenga éxito; el personal, por su parte, tiene el derecho a definir la naturaleza del programa y las prioridades de la evaluación. La distinción entre evaluación y programa como operaciones diferentes se hace confusa. Las personas del programa se convierten en coevaluadores, y los evaluadores se convierten en efecto, en asesores del programa, especialmente cuando dan al programa experiencia obtenida en la evaluación de programas similares» (Stake, 1980).

Cuando se plantea un proceso de autoevaluación se manifiesta más claramente este enfoque formativo, ya que lo que hacen los integrantes del programa es intensificar la vertiente reflexiva, dar rigor al análisis y sistematizar la vertiente crítica que casi inevitablemente acompaña a la acción. Si la evaluación es externa y está demandada por los integrantes, será fácil establecer y potenciar esta virtualidad de la evaluación. Y será más compleja, aunque posible, cuando la evaluación

sea externa y la decisión haya partido de un patrocinador con intereses (científicos o políticos) en llevarla a cabo.

La evaluación formativa respecto al programa exige, obviamente, que sea desarrollada en el curso de la acción, estableciéndose un proceso de interacción permanente entre la evaluación y la acción, proceso que tiene que estar convenientemente analizado, ya que, desde el momento que la evaluación comienza, se produce un efecto sobre el desarrollo del programa.

### **La voz de los participantes en el programa**

Los destinatarios del programa tienen en sus manos las llaves de la interpretación las claves del significado de lo que acontece. Esta valoración que los participantes realizan puede ser discrepante e, incluso, contradictoria. Por los mismos hechos y razones que unos participantes valoran positivamente un programa (caudal de conocimientos recibidos, régimen organizativo, método de trabajo empleado...) otros lo descalifican. Esa valoración es, necesaria y afortunadamente subjetiva. El evaluador no tiene que eliminar la discrepancia sino que debe explicarla, debe buscar teorías que justifiquen esas posturas divergentes.

Es lógico que existan diferentes apreciaciones en los sujetos protagonistas, pero la explicación final ha de ser coherente y aportar una interpretación de esas particulares contradicciones y de los casos atípicos o negativos (Patton, 1980). Decir que las piezas de un rompecabezas casan, no significa que todas sean iguales. Esa corroboración o coherencia estructural nos permitirá precisamente saber si existe consistencia interna en la evaluación.

La valoración que hacen los sujetos del programa es imprescindible para saber qué es lo que está sucediendo en él. De ahí que

sea necesario *dar la palabra* a los que habitualmente se quedan callados salvo para ser preguntados sobre lo que han aprendido, para que hablen sobre la valoración que hacen del programa, sobre el significado de los procesos y de las relaciones, sobre la eficacia de lo que se está haciendo, sobre los motivos que está despertando.

«Para conocer la realidad y sus significados relevantes hay que sumergirse en el curso real y vivo de los acontecimientos y conocer las interpretaciones diversas que hacen de las mismas aquellas que las viven. Esto implica que tanto las fuentes de los datos como los destinatarios de los informes son todos cuanto participan en un programa educativo y por tanto el modelo de evaluación debe ser democrático» (Pérez Gómez, 1983).

Esta exigencia no nace solamente de una concepción democrática del proceso (Mac Donald, 1976; Elliot, 1981; Stenhouse, 1982) sino de la exigencia de comprensión de una realidad cuyas claves semánticas se encuentran en poder de los propios participantes. El principio de *reflexividad* (Schwartz y Jacobs, 1984) dice que la reconstrucción de la realidad requiere el uso de unos códigos cuyos ejes están precisamente escondidos, bajo la capa superficial de los hechos, en esa misma realidad. De tal modo que para interpretarla no sólo hay que observar lo que en ella sucede sino es necesario recoger las claves de interpretación que conocen los *nativos*, es decir los miembros del programa.

Para que esa información sea rigurosa y ética es preciso crear unas estrategias de negociación y unas condiciones en el desarrollo metodológico que permitan garantizar su veracidad y su significación. Si los sujetos se ven forzados (es difícil marcar con precisión las fronteras entre coacción y persuasión) a participar en la evaluación, si no se garantiza el anonimato deseable, si no se tiene en cuenta el papel que desempeñan los informantes, si no se exigen garantías de validez a través del examen cruzado de métodos interactivos de exploración, si se solicita la información de

manera sesgada, si no se conocen y se aplican las exigencias de una exploración etnográfica, será arriesgado concluir que los participantes de la experiencia piensan algo determinado sobre ella.

### La importancia del escenario natural

Cada contexto en el que se desarrolla un programa tiene unas características peculiares que permiten conocer la realidad en sus verdaderas dimensiones.

«La forma de un programa en acción difiere de comunidad en comunidad en respuesta a las necesidades y a la tolerancia de la misma» (Weiss y Rein, 1970).

La evaluación que proponemos exige la inmersión prolongada del evaluador en el escenario natural en el que se desarrolla el programa. No es fácil siquiera imaginar una evaluación realizada por correo o bien en situaciones de laboratorio que alteran el desarrollo *normal* de la actividad.

La valoración de un programa requiere que se tenga en cuenta el contexto desde una doble perspectiva. Por un lado el *contexto diacrónico*: la consideración de los programas que se han realizado con anterioridad, los planteamientos que sobre la formación y perfeccionamiento del profesorado se ha hecho en el país o en el contexto espacial/geográfico más cercano. Y por otro lado, el *contexto sincrónico*: qué características tiene el momento, el lugar y la política general de formación y perfeccionamiento. En definitiva, es preciso definir las coordenadas estructurales, espaciales y temporales en las que se desarrollan los acontecimientos.

«El científico social no puede comprender la conducta humana sin comprender la estructura en la que los sujetos interpretan sus pensamientos, sentimientos y acciones» (Wilson, 1977).

Este enfoque de carácter etnográfico requiere una exploración asentada en princi-

pios y métodos de comprensión de una realidad desde su dinámica interna, como atinadamente apuntan Goetz y Lecompte (1988):

«El objeto de la etnografía es aportar valiosos datos descriptivos de los contextos, actividades y creencias de los participantes en los escenarios educativos. Habitualmente dichos datos corresponden a los procesos educativos tal y como éstos ocurren».

Estas mismas autoras apuntan:

«El ideal de una etnografía es describir a los participantes, el escenario, y las circunstancias con tanta claridad que la imagen reproducida componga una fotografía verbal» (Los subrayados son míos).

Los evaluadores deben ser sensibles a los matices del programa, a las peculiaridades del contexto, a la particularidad de la organización esforzándose en mirar el acontecer desde la perspectiva de los participantes. El evaluador se acerca al escenario de los hechos para preguntarse por lo que ocurre en ese contexto natural.

«¿Qué está sucediendo aquí? parece a primera vista una pregunta trivial. Pero lo más obvio, las cosas rutinarias de la vida diaria en las situaciones educativas se convierten en invisibles por ser demasiado habituales. Como dice el proverbio, "el pez es la última criatura en descubrir el agua". La investigación de campo en las situaciones educativas, a través de su inherente actividad reflexiva, ayuda a los investigadores y a los que realizan tareas prácticas en educación a "convertir lo familiar en extraño" y, por tanto, de nuevo interesante. Lo que está ocurriendo puede llegar a hacerse visible y a ser documentado sistemáticamente» (Erikson et al., 1980).

Para evaluar un programa de perfeccionamiento es necesario, pues, conocer las situaciones naturales en que se desarrolla la actividad, tener en cuenta las interacciones que se establecen entre los sujetos, conocer cuáles son las expectativas despertadas por el curso, los criterios por los que se han adscrito los profesores al mismo, quién lo ha convocado y con qué finalidad, quién y cómo va a manejar los *resultados*, qué se exige a los participantes, qué criterios de evaluación interna se ha planteado, qué tipos de tareas de aprendizaje se practican, qué estrategias de

implantación de normas se desarrollan, etc.

Hemos apuntado algunas claves *meta o supraprogramáticas* que nos permiten conocer, comprender e interpretar con más rigor lo que sucede en el programa. Los condicionantes externos de convocatoria, organización, control, coherencia institucional, etc. permitirán descifrar muchas de las claves internas del funcionamiento, del ambiente y de la eficacia del programa.

El proceso de evaluación externa puede constituir una amenaza para la dinámica espontánea y natural del programa, sobre todo cuando se aborda desde la concepción etnográfica. Ahora bien, mientras más prolongado sea el tiempo de permanencia, más amplio el ámbito de la observación y más claros los planteamientos de los evaluadores, el condicionamiento distorsionante será menor. El observador llega a ser menos extraño a medida que pasa el tiempo, su *presión* es mayor cuando actúa sobre un individuo que cuando lo hace sobre un grupo numeroso y, desde luego, es considerado de forma diferente si negocia honestamente la intervención que si aparece en la actividad sin conversaciones previas o por imposición de la autoridad.

### La interpretación desde la perspectiva etnográfica

La perspectiva cualitativa de la evaluación establece valoraciones subjetivas desde diversos frentes y momentos del análisis. Ya hemos dicho que pueden existir multiplicidad de explicaciones por parte de los participantes sobre el valor del programa. Esos juicios tienen carácter subjetivo. Pero esa condición no ha de entenderse como una limitación sino como una necesidad, ya que interesa el desarrollo de una perspectiva subjetiva sensible y autocrítica, acorde con la naturaleza de los fenómenos estudiados.

El evaluador también ha de realizar valo-

raciones subjetivas desde su particular concepción de la realidad educativa y del fenómeno que está analizando. Esta interpretación no es un proceso final que tiene lugar una vez almacenados y ordenados los datos recogidos durante la exploración, sino permanente diálogo entre los hechos y las teorías.

«Los supuestos etnógrafos también tienen la idea equivocada de creer que la descripción es un paso que debe completarse antes de que se dé el siguiente paso del análisis. En la etnografía los datos y su interpretación se desarrollan conjuntamente uno informando al otro. Los datos adicionales proporcionan una ilustración, complementan la exactitud de lo desarrollado y sugieren vías para una nueva investigación. El trabajo de campo y la interpretación marchan uno al lado del otro como pasos concurrentes más que secuenciales» (Wolcott, 1985).

El evaluador, una vez en el campo, a la luz de los hechos y de las interpretaciones de los participantes, elabora su explicación de la realidad. Esa explicación tiene en cuenta las opiniones de los sujetos, pero realiza sobre ellas (sobre cada una y sobre su capacidad conjunta de explicar estructuralmente la realidad) una metalectura rigurosa, ya que los interesados pueden hacer análisis irracionales, arrastrados por su posición en la realidad, por una argumentación defectuosa, por los intereses en juego, etc. Su percepción de la realidad puede ser errónea, ilusoria o engañosa.

«El comportamiento de los seres humanos está principalmente constituido por sus acciones, y es rasgo característico de las acciones el tener un sentido para quienes las realizan y el convertirse en inteligibles para los otros sólo por referencia al sentido que les atribuye el actor individual. Observar las acciones de una persona, por tanto, no se reduce a tomar nota de los movimientos físicos del actor sino que hace falta una interpretación por parte del observador, del sentido que el actor confiere a su conducta». (Carr y Kemmis, 1988). (El subrayado es mío).

El hecho de que el evaluador realice una interpretación subjetiva de las evidencias fá-

cticas que ha recogido en la realidad, no quiere decir que falte rigor en la evaluación cualitativa. Existen criterios de validez interna que permiten conocer la consistencia del análisis.

«Un estudio cualitativo no es un análisis impresionista, informal, basado en la mirada superficial a un escenario de personas. Es una pieza de investigación sistemática conducida con procedimientos rigurosos, aunque no necesariamente estandarizado» (Taylor y Dogdan, 1986).

### Comprensión específica en términos de detalles concretos

Responder a la pregunta sobre *cómo funciona un programa* en términos genéricos como *muy bien* o *muy mal* apenas si aporta información relevante y rigurosa sobre el mismo. Una u otra expresión pueden ser utilizadas por dos sujetos por motivos no solamente distintos sino opuestos.

El planteamiento cualitativo exige que se recojan de la realidad patrones de comportamiento y evidencias, minuciosamente descritas, que sean objeto de análisis.

«En la investigación de campo, la tarea de comprensión se lleva a cabo de modo autoconsciente y riguroso. Se pone mucho cuidado en hacer inferencias sobre lo que está pasando y sobre cuáles son los significados locales de modo que puedan apoyarse en la evidencia, reflejando las interpretaciones de la gente estudiada» (Erikson et al., 1980).

La interpretación de esas descripciones lleva al evaluador a seguir explorando la realidad de otra forma, buscando explicaciones, realizando enfoques progresivos sobre algunas parcelas de la realidad, recabando información sobre algunos aspectos, de tal manera que se produzca una *espiral de comprensión de la realidad*.

Este proceso de búsqueda y de reflexión, este ejercicio de indagación minuciosa y de interpretación elaborada, que requiere tiempo, preparación y actitud abierta para no tratar de someter la realidad a nuestros códigos

de interpretación sino a la inversa, potencian las destrezas del evaluador (y del participante en la medida que está implicado en el proceso) y permiten una mayor capacidad de comprensión. El evaluador no sólo comprende así más rigurosamente la realidad sino que, en el proceso mismo, desarrolla la sensibilidad y la capacidad para entenderla más plenamente.

No nos estamos refiriendo a una enumeración minuciosa de comportamientos o sucesos fragmentados, asépticamente descritos, sino a una rigurosa interpretación de los mismos que se abra a una teoría del comportamiento cultural. Las interpretaciones no se observan directamente, han de ser inferidas. Pero, para ello, es necesario disponer de abundantes y precisos datos recogidos en el campo de actuación.

### Métodos de carácter etnográfico

El CNS (Centro para las Nuevas Escuelas) de Chicago viene realizando desde hace años evaluaciones institucionales de naturaleza etnográfica. Los métodos de exploración que utilizan recogen información esencial y valiosa sobre la vida del Centro, sobre el contexto, sobre el reaprovechamiento formativo, el proceso de trabajo y los resultados sutiles o no pretendidos.

«Los investigadores que utilizan y dirigen evaluaciones de los programas sociales y educativos se han interesado cada vez más por los métodos antropológicos. Este interés ha estado favorecido por la insatisfacción con la información proporcionada por los métodos cuantitativos y el reconocimiento de la unidad e importancia de los datos que pueden recogerse mediante los métodos naturalistas-cualitativos» (CNS, 1976).

Los métodos que se utilizan en la evaluación cualitativa de programas son de carácter etnográfico, a saber: la observación participante y prolongada en el campo, la entrevista en profundidad a los participantes (tan-

to individual como en grupo), el análisis de documentos, el diario (de los participantes y del evaluador), etc.

«Esta perspectiva (etnográfica) requiere una transformación profunda e imaginativa en la metodología de la evaluación. Los diseños experimentales preordenados son sustituidos por diseños flexibles, emergentes y progresivos, que surgen de la dinámica de la investigación misma, en el proceso de interacción y diálogo entre el evaluador y la realidad. Por otro lado, se utilizan técnicas de investigación, como la observación participante y la entrevista etnográfica, que al contrario de los cuestionarios o las pruebas psicométricas, se muestran mucho más sensibles para captar y no tergiversar la singularidad contextual de los acontecimientos y la manera en que la gente justifica y explica sus acciones» (Angulo, 1978).

Una de las características sustanciales de la metodología de la investigación cualitativa es la multiplicidad, de tal manera que se extraigan datos de la realidad que permitan ser contrastados desde el prisma del método. Esto permitirá realizar *exámenes cruzados* de los datos conseguidos. Se puede contrastar lo que se ha observado con/en las entrevistas a los participantes, con los documentos que se han analizado, etc. Esta *circularidad y complementariedad metodológica* permite establecer procesos de exploración en espiral: los datos recogidos en la observación son objeto de trabajo en la entrevista y los resultados de ésta permiten dirigirse luego a la realidad mediante el *zoom óptico atencional*.

«Lo que al evaluador le importa es conseguir su información en cantidad suficiente de numerosas, independientes y fiables fuentes, de forma tal que llegue a calibrar eficazmente el status del programa, por complejo que éste sea» (Stake, 1980).

También se utilizan como instrumentos, en este tipo de evaluación, los cuestionarios y las escalas de estimación como instrumentos cuasi-observacionales y como formas de interrogación masiva. No se utilizan de forma exclusiva sino como cauces que permiten recoger una información de amplio espectro que puede ser sometida al contraste y la pro-

fundización a través de la entrevista y la observación en el campo. Los cuestionarios pueden construirse después de un tiempo de observación de la realidad, de tal manera que recojan la información más pertinente, necesaria y valiosa.

«Hay quienes excluyen totalmente el uso de cuestionarios con la excusa de que pertenecen a un estilo de investigación cuyos supuestos básicos son diametralmente opuestos a la etnografía. Uno de éstos es la creencia de que los hechos sociales pueden medirse de la misma manera que los hechos naturales, de donde el empleo de medidas objetivas y cuantificables tales como los cuestionarios, escalas de actitud, experimentos clínicos controlados y tests estadísticos de distribución, correlación y significación. Sin embargo, en los últimos tiempos esta posición se ha suavizado, de modo que cada vez son menos los investigadores que ven en esto una dicotomía de paradigmas y más los que lo consideran una dimensión con diferentes polos» (Woods, 1987).

En este mismo sentido se manifiestan otros autores, como Hammersley (1984), Cook y Reinhardt (1986), Davies (1984), Ball (1981), etc.

Existe un peligro en la utilización de los métodos de carácter etnográfico: su aparente sencillez. Puede entenderse que es difícil aplicar un análisis de varianza a los datos obtenidos en un diseño experimental, pero que es, por el contrario, muy fácil realizar una entrevista en profundidad o hacer una observación rigurosa de un fenómeno educativo. Una forma elemental de comprobar que esto no es así es meterse en el crisol de la experiencia. Cuando el evaluador se sumerge en el campo, percibe fácilmente la dificultad de buscar datos (observar no es sólo mirar sino indagar en la realidad) y dar interpretaciones precisas a través del establecimiento de redes de significados.

Se hace, pues, necesaria una formación específica en el uso de estas técnicas que ha de ir acompañada por el dominio conceptual de las teorías explicativas de los fenómenos escolares y de los procesos de enseñan-

za/aprendizaje. En definitiva, de todo aquello que permite una explicación coherente y profunda de la educación.

Si la evaluación se hace en equipo, como es deseable por múltiples motivos, se puede dedicar cada uno de sus miembros a trabajar más intensamente aquel método en el que se sienta más cómodo, más formado, más sagaz y más suficiente.

### El evaluador como instrumento primordial

De todo lo dicho se puede deducir la importancia que tiene el evaluador etnográfico. El es la piedra angular de este edificio indagador. Su preparación, su actitud, su habilidad para la negociación, su forma de relacionarse con los sujetos, etc. son sustanciales para la eficaz realización de los proyectos.

El evaluador no es una máquina de explorar. En la evaluación naturalista en la que existe una implicación con los sujetos, en la que el guión preordenado no se lleva a la práctica inexorablemente, en la que la agudeza de la mirada y de la mente del evaluador filtra el sentido de la realidad, en la que sus códigos de interpretación buscan el significado y dan valor a los hechos, en la que se construyen y generan teorías a través de un complejo ejercicio de reflexión, el evaluador es el eje neurálgico de la actividad. Para bien o para mal, es el evaluador el instrumento principal que capta la realidad y la analiza desde su peculiar óptica científica, ideológica y didáctica.

El conocimiento del evaluador, de los avatares que ha vivido en el decurso del trabajo (reflejados en el diario que recoge sus impresiones, conjeturas, temores, sentimientos, dificultades, problemas, implicaciones, etc.) permite al lector disponer de una fuente de validación, puesto que cada uno evalúa, en cierta medida, como sabe y como es. Davies (1982) se muestra interesado en conocer al-

gunas referencias biográficas del evaluador, de tal modo que le permitan acercarse a la identificación de esos datos personales que establecen las coordenadas de su trabajo. Dice:

«Al leer investigaciones de este tipo (de carácter cualitativo), pero sin referencias personales del autor, me he sentido más de una vez aburrido, por falta de esas referencias, y me he sorprendido volviendo una y otra vez a la sobrecubierta y a las primeras páginas del libro en busca de alguna información, aún lejana, acerca de la persona que lo había escrito».

La decisiva relación del evaluador con el programa que evalúa, con las personas que lo desarrollan, con los patrocinadores que lo impulsan o sufragan, exigen del evaluador unas capacidades que trascienden su preparación intelectual. A lo largo del proceso (desde la negociación inicial hasta la entrega del informe), el evaluador puede ser asaltado por sentimientos de escepticismo, soledad, aburrimiento, decepción, ansiedad, estancamiento, fracaso, etc. Puede también quedar prendido en las redes de la comunicación, de tal forma que se convierta en un *nativo* que alarga lo más posible la retirada del campo y que pierde la independencia de criterio que exige la evaluación externa.

«La etnografía puede ser una intensa experiencia solitaria, sobre todo si se trata del observador no participante. Uno es entonces marginal, un mirón extraño, y todas esas posiciones están cargadas de sentimientos de ansiedad y alienación de ser sólo a medias, de no estar ni de un lado ni de otro. La vida de la escuela continúa, la gente se afana en sus ocupaciones y en esta actividad el etnógrafo no pinta nada» (Woods, 1987).

En otro lugar de la misma obra escribe:

«La recogida de datos puede ser excitante y gratificante, pero también puede ser aburrida y frustrante. Por ejemplo, hay muchas lecciones que no presentan ningún interés para un espectador adulto, y es probable que la mayor parte de lo que observa en una escuela sean lugares comunes»

A lo largo del trabajo, el evaluador puede encontrarse con diversos tipos de conflictos,

sea con el patrocinador, sea con los participantes en el programa o con los miembros del Centro. En caso de existir diferentes bandos en el seno de la comunidad que se evalúa, es posible que uno de ellos intente convertir al evaluador en un aliado. De una u otra forma le presionará y le utilizará como un arma contra los otros, pretenderá brindarle una información favorable y obtener de él los datos deseados. El evaluador, que llegó a la experiencia como un agente externo, va haciéndose con una información que le hace depositario de poder. Las tensiones que se generan en torno a él, en forma de miedo, agresividad, recelo o halago, pueden incidir en su estabilidad emocional.

Uno de los modos de sentir el fracaso es *el síndrome del otro sitio*, que consiste en una sensación persistente de que las cosas importantes están sucediendo en otro lugar. Y otro, *el síndrome del vacío* que se podría describir como un conjunto de síntomas que hacen sentir al evaluador la sensación de estar rodeado de silencios, evasivas, ocultaciones, tópicos, frases hechas, etc. que le impiden tomar contacto profundo con la realidad. Es cierto que ese cerco de silencio es una parte de la realidad, pero cuando el evaluador no logra vadearlo tiene la sensación de estar perdiendo el tiempo y de ser burlado por los participantes. Este peligro es más probable cuando la evaluación externa ha sido promovida por instancias jerárquicas institucionales, sin la demanda previa de la comunidad.

Reflejar esos procesos en los informes ayuda a conocer quién es el autor del trabajo, cómo concibe la realidad, desde qué problemática la aborda, qué repercusiones han podido tener sus actitudes en la dinámica interna del programa, etc.

«El primer paso hacia una solución consiste en considerar estos problemas (se refiere a los problemas psicológicos del evaluador en contacto con su actividad) como parte de la investigación y documentarlos y analizarlos. Esto aumenta las probabi-

lidades de superarlos en vez de quedar aplastado por ellos» (Woods, 1987).

Otros evaluadores podrán verse reflejados en esos sentimientos, considerar las propias reacciones como *normales* y aprender a tratarlos de forma teórica y a responder a ellos de forma práctica. La interacción con una realidad viva y dinámica, con unas personas que están trabajando en su actividad y que pueden percibir al evaluador como un juez, un intruso o un chivato, despierta unas actitudes en el evaluador que, al ser analizadas, sirven de referencia (aquí nos encontramos con otra parcela de transferibilidad de la que luego hablaremos) para el aprendizaje del lector, que extrae sus conclusiones de los informes en una tarea de elaboración personal (readermade).

Otra forma de trabajar este tipo de situaciones del evaluador es la actuación en equipo. No es fácil que todos los miembros del equipo atreviesen por etapas idénticas y, aunque así fuera, la posibilidad de analizarlas permite un tratamiento teórico y emocional de las mismas. Esa dinámica de reflexión compartida tiene un efecto catártico respecto al individuo y una dimensión teórica que facilita la comprensión de las interacciones del evaluador con la realidad. Por otra parte, permite que se elaboren decisiones, tanto respecto a los individuos como al grupo que faciliten la superación de los problemas y la adecuación a las nuevas situaciones generadas por ellos. *La réplica paso a paso* exige que se contemplen esos procesos actitudinales que justifican algunas veces los cambios metodológicos, temporales y estratégicos del proceso de evaluación.

### El examen cruzado de la información

Una vez que se ha realizado la observación, celebrado las entrevistas, aplicado los cuestionarios, recopilado los materiales, ana-

lizado los diarios, etc. hay que acometer la compleja y difícil tarea de la interpretación y el análisis de los mismos. Teniendo en cuenta que esta no es una tarea final sino que se establece en diversas fases del proceso de evaluación. La interpretación no es una actividad terminal sino un permanente diálogo con los hechos.

Uno de los problemas con los que puede encontrarse el evaluador cualitativo es la *asfixia de los datos*. Es probable que el caudal de la información sea tan amplio que el evaluador tenga dificultades para no ser aplastado por él, entrando en una complicada actitud que le hace ampliar indefinidamente el periodo de exploración, dilatando las fases de contraste y de análisis más formales. Otro problema con el que se encontrará, casi inevitablemente, es el de una *heterogeneidad de la información* que puede tener visos de caótica.

Los datos de fuente única tienen menos validez que los que han pasado por el tamiz de la contrastación. Una de las formas de realizar el examen cruzado es el proceso singular o múltiple de la triangulación.

Se entiende por triangulación, según la definición clásica de Denzin (1978) «la combinación de metodologías en el estudio de un mismo fenómeno».

La triangulación permite contrastar los datos pero también es un modo de obtener otros que no han sido aportados por un primer nivel de lectura de la realidad, a través de la capa más superficial del acontecer.

«El principio básico subyacente en la idea de triangulación es el de recoger observaciones/apreciaciones de una situación o algún aspecto de ella desde una variedad de ángulos o perspectivas y después compararlas y contrastarlas» (Elliot, 1980).

Vamos a hacer referencia a cuatro tipos de triangulación que se pueden llevar a cabo en la evaluación de programas de perfeccionamiento, aunque no es preciso utilizarlas todas respecto a los mismos datos o procesos:

*Triangulación de métodos:* El evaluador del programa contrasta la información obtenida a través de tres instrumentos de exploración. Por ejemplo, la de la observación, la de la entrevista y la de los cuestionarios. En este contraste aparecerán, probablemente, discrepancias y coincidencias sobre la descripción, interpretación y valoración de los mismos hechos. El evaluador se preguntará de dónde proceden, por qué se producen y qué valor tienen. Algunas de ellas podrán ser explicadas por la naturaleza del método a través del cual se ha recogido la información. Si los participantes niegan en las entrevistas los hechos que se han recogido fidedigna y repentinamente en la observación por varios evaluadores, habrá que pensar que hay motivos para que los protagonistas intenten negar la realidad... El contraste permitirá determinar, a través del análisis riguroso, explicarlas y decantar la interpretación final hacia una síntesis integradora. De ese proceso de triangulación saldrán nuevos interrogantes sobre la realidad que dirigirán las sucesivas exploraciones.

*Triangulación de sujetos:* Se trata de contrastar los puntos de vista de los participantes en el programa. Los implicados en la actividad tienen diferentes criterios y posturas para interpretarla. Algunos nacen de la cosmovisión de cada sujeto. Otros de papel que desempeñan en la actividad. No tienen probablemente el mismo enfoque los responsables del programa que los participantes en el mismo.

«Es deseable entablar discusiones sobre puntos de desacuerdo entre varias partes implicadas, preferiblemente bajo mediación de una parte neutral». (Elliot, 1980).

Leiter (1986) llama validez cara a cara ("face validity") a la triangulación de sujetos porque nos permite contrastar adecuadamente los diferentes puntos de vista sobre una misma realidad.

La triangulación de personas puede reali-

zarse directamente (por ejemplo, evaluador / responsables del programa / participantes) o bien de forma escalonada (evaluador / responsables, evaluador / participantes. . .). Charles Hall (1986) propone, en un ingenioso artículo titulado «Cómo lograr la triangulación cuando sólo hay dos en el cuadrilátero» una forma de triangular la información a través de la grabación de vídeo en el caso de que no se disponga de un observador externo. Hacemos esta sugerencia para invitar al evaluador a explorar diferentes estrategias de contrastación de datos, sin limitarse a las que se proponen como referencia.

Las sesiones de trabajo deben prepararse convenientemente, a través del estudio de la información previa. A ella se pueden llevar los informes y los diarios de los participantes y del evaluador, con el fin de disponer de elementos nucleares de diálogo.

*Triangulación de momentos:* Un programa de formación se desarrolla en una diacronía que está marcada con hitos factuales. En el *antes* del programa se sitúa la iniciativa y el diseño del mismo. En el *durante* se escalonan temporalmente diferentes fases de acción y en el *después* se producen algunos efectos y las repercusiones producidas en los sujetos y en los ambientes de origen (o nuevos ambientes).

La triangulación temporal recoge información contrastada de los diferentes momentos, utilizando los tres momentos como vértices de la reflexión. Ese tratamiento que tiene como punto de referencia el tiempo puede aplicarse a una parte del programa, a uno de sus elementos, a una de sus fases. Antes de producirse la sesión el evaluador pregunta a los interesados qué es lo que esperan de ella, cómo la están preparando, qué actitudes albergan en la prospección del hecho educativo... Durante el proceso, observa el desarrollo de la actividad y, después de finalizada, vuelve a preguntar a los participantes, ahora sobre su visión retrospectiva de la se-

sión y sobre la valoración que hacen de la misma.

*Triangulación de expertos:* El evaluador se convierte a lo largo del proceso en un participante en la acción. La llega a contemplar desde su interior. De ahí que pueda efectuar otro tipo de triangulación sobre los informes obtenidos con expertos que no han participado en el proceso de evaluación. Se produce así una nueva espiral de comprensión, ya que se somete a consideración el contenido de la información y también los aspectos metodológicos y los cambios efectuados en el diseño de la exploración.

De esa triangulación pueden derivarse modificaciones en el planteamiento de la evaluación y, desde luego, pueden analizarse los datos desde la perspectiva complementaria del que no está implicado en un proceso, en un papel y en una función social, psicológica, metodológica y éticamente comprometida.

### La interpretación de los datos

En este modelo de evaluación no se produce una interpretación estadística de los resultados a través de estudios de significatividad estadística que confirman o rechazan las hipótesis nulas formuladas al inicio.

La recogida de datos permite ir haciendo una interpretación tentativa que va luego desarrollándose progresivamente a medida que avanza el proceso de evaluación. Pero llega un momento en que hay que dar cuerpo a un caudal disperso y quizás caótico de datos y de interpretaciones parciales y provisionales.

«Llega un momento en que la masa de los datos incorporada a las notas de campo, transcripciones, documentos, ha de ser ordenada con una cierta sistematicidad, en general mediante la clasificación y la categorización. En un nivel elemental esto se aplica simplemente a los datos propios. En esta etapa puede que no haya formación de conceptos, importación o descubrimiento de teorías, creación de nuevos pensamientos. El objetivo es dar al material una forma que conduzca a tales fines, y esto significa or-

denar los datos de una manera coherente, completa, lógica y sucinta» (Woods, 1987).

Lacey (1976) habla de una *espiral de comprensión* que exige recorrer caminos de ida y vuelta entre el método y la realidad. Está siempre al acecho el peligro de la simplificación y de la atomización en la interpretación de una realidad compleja. Los cauces de significado se pueden ir produciendo en cascada, depurándose progresivamente en una conceptualización de progresivo carácter sintético. La densidad de los diferentes niveles de análisis no viene dada por la progresiva complejidad de los contenidos sino por el nivel de abstracción de los criterios de análisis.

Los códigos de interpretación de los que parte el evaluador no han de ser determinantes. De lo contrario, forzaría la interpretación para que se confirmasen sus planteamientos teóricos, obligaría a hablar a la realidad en el lenguaje que él entiende y que le gusta oír. Taylor y Bogdan (1986) apuntan atinadamente:

«La regla cardinal de la codificación en el análisis consiste en hacer que los códigos se ajusten a los datos y no a la inversa».

Peter Woods (1986) plantea un sistema de análisis arbóreo presidido por el eje conceptual. El conjunto de los datos, considerado como una totalidad proveniente de todo el proceso de exploración, es sometido a un creciente estudio que pasa por distintos niveles caracterizados por una progresiva complejidad conceptual estructuradora. Se plantea en primer lugar un análisis especulativo consistente en una reflexión tentativa que se produce en el momento de hacer el registro. Luego se realiza una tarea de clasificación y categorización consistente en «identificar las categorías más importantes que, a su vez, pueden dividirse en grupos. Los datos pueden clasificarse de acuerdo con estas categorías, cuya naturaleza dependerá del tipo de estudio y de los intereses en juego».

La formación de conceptos es una estrategia de análisis que permite descodificar fragmentos de hechos con propiedades estructurales. Los conceptos no vienen expresados en la capa superficial del desarrollo del programa sino que son formulados por el evaluador, utilizando a veces términos del propio grupo analizado. Los conceptos adoptan, en ocasiones, la fórmula de «símbolos culturales, normalmente codificados en términos nativos» (Spradley, 1979). Los modelos encierran una importante capacidad de representación de la realidad. Los procesos complejos que se desarrollan en determinados programas de perfeccionamiento pueden comprenderse más fácilmente utilizando modelos representacionales. Y así, por ejemplo, podremos hablar de un modelo de formulación asentado en la racionalidad teórica y de otro que descansa en la racionalidad práctica (Pérez Gómez, 1987). El peligro que encierran los modelos es el de encorsetar la realidad encerrándola en una descripción teórica, ya que no existe ningún modelo teórico al que se ciña fielmente la realidad.

Finalmente, se podría formular o reformular una *teoría*, como conocimiento especulativo que se puede considerar independiente de toda aplicación. La teoría sirve para explicar las relaciones existentes entre los fenómenos, de forma que ayuda a la comprensión de la realidad. El evaluador puede utilizar teorías para el análisis, pero también ha de estar atento a su reformulación, remodelación, comprobación y creación. Así se podrá avanzar el conocimiento sobre la realidad. De esta manera se establece un proceso simbiótico entre teorías existentes y aplicadas y teorías en reformulación y creación. Esta dialéctica hace progresar la ciencia y brinda nuevas explicaciones, más profundas, más ricas, más variadas de la realidad. Las teorías no brotan perfectamente elaboradas sino que van *construyéndose progresivamente*.

El carácter descriptivo de la etnografía

puede conducir al evaluador a un rechazo del marco teórico de referencia. Sin embargo, la teoría es una gran ayuda para la comprensión de la realidad, con tal de que no se pretenda forzar la interpretación para confirmar y salvaguardar las *concepciones teóricas apriorísticas* del evaluador o del patrocinador.

La construcción de la teoría supone un salto conceptual de la mera descripción a la articulación explicativa que permite conocer las leyes por las que se rige la realidad educativa. Hace falta, pues, una buena dosis de imaginación y de creatividad, juntamente con el rigor metodológico de la observación, el registro y la interpretación.

#### El informe y la retroalimentación

El proceso de escritura ha de ser un instrumento de producción de ideas, no sólo un medio de comunicación de las informaciones. Un mal informe invalida un buen proceso de evaluación. De ahí la necesidad de desarrollar la capacidad de transcribir fielmente los registros y las interpretaciones. La redacción del informe requiere unas destrezas de descripción fiel y amena de la realidad. La capacidad de los hechos pueden esterilizarse ante la incapacidad narrativa.

Resulta esencial *garantizar la libertad y la independencia del evaluador* a la hora de hacer el informe. Las presiones de carácter psicológico, político, ideológico y social, pueden desvirtuar la naturaleza y los resultados de la evaluación. Hay que proteger al evaluador, hay que fortalecer su papel con el fin de dar garantías de la honestidad y el rigor del informe.

El evaluador ha de ser claro y preciso. Para ello ha de utilizar un lenguaje asequible al destinatario y que, a su vez, sea el reflejo del modo que tienen de valorar los programas quienes participan en ellos. Un lenguaje arcano, cargado de tecnicismos, impide que sea

comprendido por muchas personas. Es deseable utilizar un lenguaje natural, alejado de los términos y los giros crípticos con que muchas veces se reviste el lenguaje científico. Es un error pensar que por ello el trabajo pierde rigor y seriedad.

El rigor de las interpretaciones ha de estar basado en la riqueza y precisión de la indagación. Por eso, las conclusiones han de estar avaladas por la descripción metodológica que ha precedido a su elaboración. De ahí que sea aconsejable disponer del material documental de referencia que ha servido de base al trabajo, presentado en forma de apéndice o dossier complementario. Schwartz (1984) habla del método de las *confesiones verdaderas*, según el cual habría que expresar cómo y por qué se condujo el autor, aunque advierte que esa pretendida fidelidad

«...no resuelve el problema de que nuestros métodos de observación y análisis de la realidad pueden no funcionar en el mundo social».

Cuando hablamos de conclusiones en el informe no lo hacemos en el mismo sentido que las formulan los trabajos de carácter estadístico. En la evaluación etnográfica son los mismos lectores los que formulan sus propias conclusiones, deduciendo su interpretación de las descripciones y de los comentarios que el evaluador realiza sobre ellas en el informe.

Una vez analizado el informe por sus destinatarios (hay que precisar en la negociación inicial cómo, cuándo y a quién va a llegar ese informe) debe ser estudiado con ellos, de modo que permita abrir un proceso de retroalimentación y de mejora de la profesionalidad. En esta tesis insiste Helen Simons (1987) cuando habla del proceso de autoevaluación de las escuelas. La evaluación encierra dentro de sí misma unos procesos de análisis y de crítica que pueden convertirse en un instrumento de cambio. Pero, es la reflexión que nace de la discusión del informe lo que permite incidir más abierta y fundadamente en

la transformación de la realidad.

El proceso de retroalimentación ha de estar libre de peligros que disminuyan la eficacia: el retraso indefinido en la entrega y discusión, la descalificación de resultados rigurosos porque no resultan gratos o la potenciación de algunas conclusiones débiles porque privilegian intereses personales o institucionales...

Hacemos aquí una llamada sobre el tema de la *ética de la evaluación cualitativa*, ya que una de sus vertientes es la utilización caprichosa, arbitraria o interesada por parte del patrocinador de los informes de la evaluación. El patrocinador convierte la evaluación en un apoyo técnico para toma de decisiones políticas de dudosa lógica y de difícil justificación. Un informe permite suprimir un programa, retirar un presupuesto, apartar a una determinada persona...

La evaluación no es un proceso que se cierra sobre sí mismo. Tiene por finalidad conocer la realidad, comprenderla fielmente, ofrecer referencias rigurosas para cambiarla y alumbrar nuevas teorías para explicarla más profundamente. Los cambios podrán ser evaluados en un nuevo proceso de reflexión sistemática...

Ese planteamiento se subraya en la consideración de la evaluación curricular como una estrategia de conseguir el perfeccionamiento del profesorado. (Sancho Gil, 1987).

#### Metaevaluación: las fuentes de la credibilidad

La metaevaluación es un proceso de análisis que consiste en hacer un juicio de valor sobre la misma evaluación. La lectura de un informe sobre la evaluación de un programa de perfeccionamiento merece la credibilidad del lector si se ajusta a unos criterios rigurosos que garantizan su validez.

«Mientras los racionalistas están interesados en

proteger la investigación contra factores de competencia o encubrimiento (fuentes de error), que se dice confunden la investigación, los naturalistas desean tomar en consideración el desconcertante conjunto de factores interrelacionados en las muestras que han de afrontar y que dificultan enormemente la interpretación. La solución al problema que dan los racionalistas consiste en abstraer variables de interés especial, controlando o aleatorizando el resto. La solución naturalista es tratar con las muestras en su integridad y llevar a cabo ciertas acciones que den cuenta de las complejidades» (Guba, 1981).

Podríamos hablar de *validez interna* cuando el proceso de análisis refleja con precisión lo que sucede en el programa, las interpretaciones se derivan rigurosamente de los hechos y cuando los cambios que señala como deseables constituyen una mejora para el mismo. Esa validez se consigue *con el trabajo prolongado en el campo, la separación periódica de establecimientos de pistas de reflexión, la corroboración o coherencia estructural, la recogida de material de adecuación referente, la justificación de los cambios metodológicos realizados durante el proceso de evaluación, la réplica paso a paso, el rigor interno en la aplicación de cada método, el juicio crítico de expertos, etc.*

La *validez externa* ofrece garantías de rigor al establecer criterios que permiten transferir los resultados a otras situaciones similares y elevar los hallazgos conseguidos en la evaluación de un programa concreto, tema que ahora nos ocupa, a la categoría de aportación teórica para la comprensión del perfeccionamiento.

La evaluación naturalista no formula generalizaciones del tipo de las que surgen de las investigaciones racionalistas, articuladas sobre afirmaciones estocásticas asentadas en la significatividad.

«La generalización en los estudios interpretativos... descansa en la asunción de la intersubjetividad de las interpretaciones y en la habilidad del lector de generalizar personalmente a su propia situación, más que en la generación de una investigación a la población más amplia que la muestra usada en un estudio particular» (Mc Cutcheon, 1981).

Walker (1980) habla de *transferibilidad* cuando se refiere al proceso que nos permite aplicar los hallazgos conseguidos en un caso concreto a otra situación diferente/similar. Cuando nos enfrentamos a esa nueva situación, en parte distinta (ya que no es una repetición exacta de la misma), en parte semejante (ya que tiene elementos comunes), nos preguntamos: *¿Hasta qué punto son válidos en esta nueva situación los resultados obtenidos en la primera? ¿Son aplicables a este contexto aquellas conclusiones? ¿Se producen aquí los mismos procesos? ¿Pueden ser explicados con los mismos presupuestos teóricos?...*

Una buena parte de la validez externa descansa, pues, en el análisis del lector que, una vez conocida la evaluación y la realidad nueva, establece las oportunas conexiones entre ambas. En este sentido se manifiestan Guba y Lincoln (1982):

«Si el término generalización tiene que tener algún significado, es con respecto a audiencias particulares. Es responsabilidad de cada audiencia determinar... lo que la información significa y determinar por sí misma su aplicabilidad».

Algunos criterios para conseguir la validez externa son la similitud de los contextos, la existencia de abundantes descriptivos, las afinidades multidimensionales entre las experiencias (sociales, económicas, legales...), el desarrollo y verificación correcta de la teoría, la multiplicación de evaluaciones, el análisis progresivo, etc.

### Consideraciones finales

No quiero terminar sin hacer alguna referencia a los problemas que aparecen en el horizonte a la hora de llevar a la práctica un tipo de evaluación de esta naturaleza. (Pienso en procesos de evaluación externa, alentados por la misma Administración, por la institución educativa o por determinados patrocini-

nadores. Sin dejar de lado los procesos de autoevaluación, tan deseables desde el punto de vista del perfeccionamiento).

Uno de ellos es el del tiempo que requiere y la preocupación por el fenómeno educativo que supone. Baste al efecto considerar cuánto tiempo, esfuerzo y dinero se dedica a procesos de evaluación frente al destinado a la realización de los programas.

«Los impedimentos de la autoevaluación en las escuelas citados frecuentemente por los profesores son tiempo y motivación. La mayoría de los profesores piensan que no tienen suficiente tiempo para comprometerse con esta actividad adicional y muy pocos son los que están preparados para darle prioridad, especialmente cuando, como sucede a menudo, no hay compromiso firme por parte de la jerarquía escolar de considerar los resultados» (Simons, 1981).

Otro problema es la resistencia a la consi-

deración del hecho educativo como un fenómeno institucional, que supera las fronteras del aula y rompe la concepción individualista del proceso de enseñanza/aprendizaje. (Santos Guerra, 1987). De ahí que la evaluación se plantee habitualmente como un proceso de rendición de cuentas de *cada individuo* y no como un análisis sistemático del funcionamiento de las instituciones, de los Centros, de los programas, etc. Esa concepción individualista es uno de los elementos que genera en el profesor el deseo de no ser sujeto de ningún tipo de evaluaciones...

Es preciso también contar con profesionales competentes que dominen el marco teórico y metodológico y que dispongan de tiempo destinado a una tarea tan necesaria como apasionante.

## REFERENCIAS

- ANGULO RASCO, F. (1988). Evaluación de programas sociales: De la eficacia a la democracia. Universidad de Málaga. Multicopiado.
- BALL, S. J. (1981). *Beachside Comprehensive*, Cambridge University Press. Cambridge.
- CARR, W. C. y KEMMIS, S. (1988). *Teoría crítica de la enseñanza. La investigación en la formación del profesorado*. Ed. Martínez Roca. Barcelona.
- CNS (1976). La evaluación etnográfica en educación. *Journal of Research and Development in Education*, Vol. X, nº 4.
- COOK, T. D. y REINCHARDT, CH. S. (1982). *Métodos cualitativos y cuantitativos en investigación educativa*. Ed. Narcea. Madrid.
- DAVIES, L. (1982). *Life in the classroom and Playground*. Roulledge and Keagan Paul. London.
- DAVIES, L. (1984). *Pupil Power. Deviance and Gender in School*. Falmer Press. Lewes.
- DENZIN, N. K. (1978). *The Research Act: A theoretical Introduction to Sociological Methods*. McGraw-Hill Book Company. New York.
- ELLIOTT, J. (1980). Accountability, Progressive Education and School-based Education, Cambridge Institute Education. Multicopiado.
- ELLIOT, J. (1981). *Educational Accountability and Evaluation or Teaching*. Cambridge Institute of Evaluation.
- ERIKSON, F. et al. (1980). Investigación en el campo de la educación. *Educational Research*. East Lansing. Michigan State University.
- GIMENO SACRISTAN, J. (1984). *La pedagogía por objetivos. Obsesión por la eficiencia*. Ed. Morata. Madrid.
- GOETZ, J. P. y LECOMPTE, M. D. (1988). *Etnografía y diseño cualitativo en investigación educativa*. Ed. Morata. Madrid.
- GUBA, E. G. (1981). Criterios de credibilidad en la investigación naturalista, en Gimeno Sacristán, J. y Pérez Gómez, A. (1983). *La enseñanza, su teoría y su práctica*. Ed. Akal. Madrid.
- GUBA, E. G. y LINCOLN, Y. (1982). Epistemological and Methodological Bases of Naturalistic Inquiry. *Educational Communication and Technology Journal*. Vol. XXX, nº 4.

- HAMMERSLEY, M. (1984). The researcher exposed: A natural history, en BURGESS, R.G. (Comp.): *The Research Process in Educational Setting. Ten Case Studies*. Falmer Press. Lewes.
- HULL, C.H. (1986). Cómo lograr la triangulación cuando sólo hay dos en el cuadrilátero, en Elliott, J. et al.: *Investigación/acción en el aula*. Generalitat Valenciana. Conselleria de Cultura Valencia.
- LACEY, C. (1976). Problems of sociological fieldwork: A review of the methodology of "Hightown Grammar", en Hammersley, M. y Woods, P. (Comp.): *The Process of Schooling*. Routledge and Keagan Paul. London.
- LEITER, K. (1986). *A Primer on Ethnomethodology*. Oxford University Press. New York-Oxford.
- MCCUTCHEON, G. (1981). On the Interpretation of Classroom Observations. *Educational Research*. Vol. X, nº5.
- MCDONALD, B. (1976). Evaluation and the Control of Education, en Tawney, D. (1981): *Interviewing in Case Study Evaluation*. Los Angeles. Area.
- PATTON, M. Q. (1980). *Qualitative Evaluation Methods*. Beverly Hills. Calif: Sage.
- PEREZ GOMEZ, A. (1983). Modelos contemporáneos de evaluación, en Gimeno Sacristán, J y Pérez Gómez, A (1983). *La enseñanza, su teoría y su práctica*. Ed. Akal. Madrid.
- PEREZ GOMEZ, A. (1987). El pensamiento práctico del profesor; implicaciones en la formación del profesorado. Congreso Mundial Vasco, Bilbao.
- SANCHO GIL, J. M. (1987). *La evaluación del currículum y el perfeccionamiento del profesorado*. Universidad Central de Barcelona. Tesis Doctoral.
- SANTOS GUERRA, M. A. (1987). Organización escolar e investigación educativa. *Investigación en la Escuela*, nº 2. Sevilla.
- SIEDMAN, E. (1977). *Why not Qualitative Analysis?* Public Management Forum. Julio/Agosto.
- SIMONS, H. (1981). El proceso de autoevaluación en las escuelas, Institute of Education. Falmer Press. London.
- SIMONS, H. (1981). *Getting to know Schools in a democracy. The Politics and Process of Evaluation*. Falmer Press. London.
- SPRADLEY, J. P. (1979). *The Ethnographic Interview*. Holt. Rinehart and Winston. New York.
- STAKE, R.E. (1978). The Case Study Method in Social Inquiry. *Educational Research*. Vol. III, nº2.
- STAKE, R. E. (1980). La evaluación de programas; en especial la evaluación de réplica, en Dockrel, W.B. y Hamilton, D. (1980). *Nuevas reflexiones sobre la investigación cualitativa*. Ed. Narcea. Madrid.
- STENHOUSE, L. (1982). the Teacher as a Focus of Research and Development. East Anglia University. Norwich. Multicopiado.
- TAYLOR, S. J. y BOGDAN, R. (1986). *Introducción a los métodos cualitativos de investigación*. Ed. Paidós/Studio. Buenos Aires.
- WALKER, R. (1980). La realización de estudios de casos en Educación. Ética, Teoría y Procedimientos, en *Nuevas reflexiones sobre la Investigación Educativa*. Ed. Narcea. Madrid
- WEISS, R. S. y REIN, M. (1970). The Evaluation of broad-aim programs: Difficulties in experimental design and an alternative, En Weiss, C.H. *Evaluating Social Action Programs: Reading in Social Action and Evaluation*. Allyn and Bacon. Boston.
- WILSON, S. T. (1977). The use of Ethnographic Techniques in Educational Research. *Review of Educational Research*. Vol. XLVII, nº1.
- WOLCOTT, H. F. (1985). Sobre la intención etnográfica. *Educational Administration Quarterly*. Vol. XXI, nº3.
- WOODS, P. (1987). *La escuela por dentro. La etnografía en la investigación educativa*. Ed. Paidós/MEC. Madrid.

**SUMMARY**

*The author sets up the need of doing the evaluation of the educational institutional teachers improvement processes not as an ending appendix of the activity, but as a rigorous assesment analysis during the activity development, what makes easy the decision-making, and bring about a deeper comprehension of the phenomena.*

*He focus the attention in the qualitative evaluation as a requirement of the complex, dynamic and absolutely special nature (so that essential) and full of values of the educational phenomenon which takes place in the planns and the centers for teachers training.*

*The methodological issues coherent with the approach philosophy are also exposed , and the tecnic, social, politic and etic conditions that derives from the naturalistic idea of evaluation are established.*

*At the end, the author stops to consider the credibility sources to take into account when the meta evaluation procces is planed.*

**RÉSUMÉ**

*L'auteur établit la nécessité de réaliser l'évaluation des processus éducatifs institutionnels de perfectionnement des enseignants non comme une appendice finale de l'activité, mais comme un analyse valoratif rigoureux au cours vital de l'activité, mais comme un analyse valoratif rigoureux au cours vital de l'activité, en facilitant la prise des décisions et en produisant une compréhension plus grande des phénomènes.*

*Il centre l'attention sur l'évaluation cualitativa comme une exigence de la nature complexe, dynamique, très péculière (et, pour ça, imprescindible), chargée de valeurs du système éducatif, qui existe dans les programmes et dans les Centres de perfeccionement des enseignants.*

*On suit les questionnements méthodologiques cohérents avec la philosophie de l'approche et on établit les conditions techniques, sociales, politiques et éthiques qui se dérivent de la conception naturaliste de l'évaluation.*

*Finalment, l'auteur considère les conditions de crédibilité du processus de metaévaluation.*