

Desde [http://www.redes-cepalcala.org/inspector/DOCUMENTOS%20Y%20LIBROS/EVALUACION/EVALUACION%20Y%20CALIDAD%20\(UNED\)/UNED\\_6/webdoc6.htm](http://www.redes-cepalcala.org/inspector/DOCUMENTOS%20Y%20LIBROS/EVALUACION/EVALUACION%20Y%20CALIDAD%20(UNED)/UNED_6/webdoc6.htm)

# EVALUACIÓN DE PROGRAMAS EDUCATIVOS: MODELOS, TÉCNICAS E INSTRUMENTOS, ANÁLISIS DE DATOS Y ELABORACIÓN DE INFORMES

Por *José Luis García Llamas* (Universidad Nacional de Educación a Distancia)

Regresar a [Índice de materias](#)

---

## Contenido

1. [INTRODUCCIÓN \(ASPECTOS GENERALES\)](#)
2. [MODELOS DE EVALUACIÓN DE PROGRAMAS:](#)
3. [TÉCNICAS E INSTRUMENTOS PARA LA EVALUACIÓN DE PROGRAMAS:](#)
4. [EL ANÁLISIS DE DATOS Y LA ELABORACIÓN DE INFORMES EN LA EVALUACIÓN DE PROGRAMAS EDUCATIVOS:](#)

## [BIBLIOGRAFÍA](#)

---

## I. INTRODUCCIÓN (ASPECTOS GENERALES)

La evaluación como área de conocimiento reciente y en proceso de formación, tiene su desarrollo histórico ubicado en los Estados Unidos, con la finalidad de valorar los resultados de programas especiales diseñados para alumnos con problemas sociales, así como para asegurar que las innovaciones educativas mejoraban realmente la enseñanza. Nace pues la evaluación vinculada a la tradición positivista de la investigación educativa. De ahí que, en ocasiones, aparezcan ambos términos entrelazados y en algunos momentos confundidos en los textos.

Mientras la **investigación** se caracteriza por las notas de: probar, extraer conclusiones, no implicación del investigador, verdad científica, generalización; la **evaluación** presenta como características: mejorar, tomar decisiones, implicación del evaluador, valoración del fenómeno, sin pretensiones de generalización.

De la confluencia de ambos términos surge la **investigación evaluativa** que en palabras de DE LA ORDEN (1985) puede definirse como "el proceso sistemático de recogida y análisis de información fiable y válida para tomar decisiones sobre un programa educativo". Por su parte ALVIRA MARTÍN (1985) la define como "la acumulación de información sobre una intervención, programa o sobre su funcionamiento y sobre sus efectos y consecuencias". SUCHMAN (Cit. Stufflebeam, 1989) nos dice que la investigación evaluativa consiste en "los procedimientos de recopilación y análisis de datos que aumentan la posibilidad de demostrar, más que de asegurar, el valor de alguna actividad social". Para PATTON (1990) "el propósito de la investigación evaluativa es informar a la acción, apoyar la toma de decisiones y aplicar el conocimiento para resolver problemas humanos y sociales".

Ahora bien, existen muchas reticencias por parte de la comunidad científica de otorgar el rango de investigación a la evaluación. Sin embargo, el desarrollo de la evaluación, en su sentido más amplio, durante las últimas décadas, ha contribuido de forma decisiva a su implantación en distintas esferas de la actividad educativa y no solamente como verificación de los logros alcanzados (modelos tylerianos), sino como comprobación de los procesos seguidos (Scriven, Stufflebeam, Pérez Juste).

Dentro del marco de referencia educativa un programa, podemos entenderlo según Pérez Juste como un plan sistemático de intervención, diseñado y elaborado intencionalmente para la consecución de unos objetivos de mejora. Ello supone que todo programa educativo debe estar incardinado con los proyectos educativos de los centros y contar con la colaboración del equipo de profesores que hagan posible su procesualidad, mantenimiento o transferencia según los casos.

Conscientes de las limitaciones y dificultades inherentes a la evaluación, hemos de reconocer su presencia continuada en todo diseño de intervención educativa, desde el momento en que se identifican las necesidades, se formulan los objetivos, se desarrolla el proceso de ejecución y se analizan los resultados alcanzados. Por ello, podríamos afirmar que la evaluación de programas se configura como un componente intrínseco de los procesos sociales y educativos desde una consideración globalizada de los mismos. Por tanto, no se trata de algo puntual, sino de un *proceso* que nos permite obtener información necesaria que sea relevante, fiable, válida, adecuada y recogida a tiempo, con ello podremos emitir *juicios de valor* tanto sobre el proceso como el producto, lo que nos conducirá a la toma de decisiones de mejora. En consecuencia la evaluación no debe ser un fin en sí misma, sino que debe estar orientada al futuro. Entendemos la evaluación GARCÍA LLAMAS (1986) como "el proceso de identificación, recolección y tratamiento de datos para obtener una información que justifique una determinada decisión. Debe servir, no solo para analizar un determinado programa, sino también para ayudarnos a comprender el proceso de aprendizaje".

## **1.1. Concepto de evaluación de programas:**

Una de las definiciones clásicas en la evaluación de programas se debe a TYLER (1950) quién nos define la evaluación como "el proceso para determinar en qué medida los objetivos educativos habían sido alcanzados mediante los programas de currículos y enseñanza". En esta línea centrada en los logros WEIS (1991) nos dice que el objeto consiste en "medir los efectos de un programa por comparación con las metas que se propuso alcanzar, a fin de contribuir a la toma de decisiones sobre el programa y para mejorar la programación futura".

No obstante, las definiciones posteriores a Tyler, hacen alusión a los procesos, necesidades, contextos, inputs. Así STUFFLEBEAM (1989) nos dice que "la evaluación de programas es el proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y el mérito de las metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados".

ROSSI y FREEMAN (1991) definen la evaluación como "la aplicación sistemática de los procedimientos de investigación social en la valoración de la conceptualización y diseño de un programa, de la implementación y de la utilidad de los programas de intervención social"

PÉREZ JUSTE (1992) nos ofrece una definición muy completa sobre la evaluación de programas al afirmar que consiste en un "proceso sistemático, diseñado intencional y técnicamente, de recogida de información rigurosa -valiosa, válida y fiable-, orientado a valorar la calidad de un programa, como base para la posterior toma de decisiones de mejora tanto del programa como del personal implicado y, de modo indirecto, del cuerpo social en que se encuentra inmerso".

Del análisis de las definiciones anteriores parece desprenderse la idea de que la evaluación de programas no configura un ente aislado del resto de la actividad educativa, sino que, por el contrario, se encuentra íntimamente ligada a su desarrollo, dado que nos ofrece información sobre la marcha de los procesos sociales y/o educativos, así como sobre la valía de las propuestas innovadoras en el campo. Todo ello contribuirá de forma decisiva para la toma de decisiones que permitan una mejora en tres puntos básicos: el programa en sí mismo, los procesos seguidos y los resultados derivados de su implementación, es decir, del personal implicado en el mismo.

La evaluación de programas cumple los siguientes cometidos: diagnóstico, formativo y orientador, predictivo y de control del proceso educativo. De una parte ayuda a desvelar los puntos fuertes o débiles del programa y de otra orienta hacia el desarrollo futuro, de forma que las personas implicadas adquieran confianza en su tarea y se puedan remediar las deficiencias que se detecten.

A medida que se va desarrollando la reflexión y la acción evaluativa, se ha ido acumulando una gran cantidad de conocimientos que tienden a fundamentar teóricamente la evaluación de programas, de la que van surgiendo distintos modelos que son el fruto de las investigaciones más recientes, cuya finalidad consiste en ir creando una teoría sobre la evaluación y una gestión eficaz sobre su sistematización, ello garantizará la lógica del

procedimiento y su justificación.

## 1.2. Clasificación de los modelos de evaluación de programas:

Los modelos de evaluación admiten diversas clasificaciones. Para POPHAM (1980) existen cuatro modelos:

- \* De consecución de metas: Tyler, Hammond, Metfessel.
- \* De enjuiciamiento sobre criterios intrínsecos: acreditación.
- \* De enjuiciamiento sobre criterios extrínsecos: Stake, Scriven.
- \* De facilitación de decisiones: Stufflebeam, Guba, Alkin.

Por su parte STUFFLEBEAM y SHINKFIELD (1989) distinguen tres apartados:

1. Pseudoevaluación, dónde se incluyen las investigaciones encubiertas y los estudios basados en las relaciones públicas.

2. Cuasievaluaciones, a su vez clasificadas en:

- Estudios basados en objetivos: Tyler, Bloom, Popham.
- Estudios basados en la experimentación: Campbell y Stanley, Suchman, Cronbach, Snow.

3. Verdaderas evaluaciones, dónde se incluyen:

- Estudios de orientación de la decisión: Cronbach, Stufflebeam, Alkin, Guba, Taylor.
- Estudios centrados en el cliente: Stake, MacDonald.
- Estudios políticos: Coleman, Clark, Wolf, Weiss.
- Estudios basados en el consumidor: Scriven, Glass.

Cada uno de los modelos de evaluación de programas tiene sus virtualidades y sus inconvenientes, si bien, los más recientes prestan una especial atención a los momentos inicial y procesual frente a los meramente medicionales de comprobación de los logros alcanzados. De esta forma se presta gran atención al momento inicial de la puesta en marcha de un programa, lo que Pérez Juste define como **el programa en sí mismo**, dónde se pueden incluir la calidad del propio programa y la adecuación al punto de partida y al contexto (Stufflebeam). Un segundo momento sería **la puesta en marcha del programa**, es decir en su desarrollo, incluyendo la evaluación procesual o formativa (Scriven), lo que conducirá a la facilitación de la toma de decisiones a tiempo por parte del personal implicado. El tercer momento se centra en **los logros alcanzados por el programa**, es

decir, el producto (Stufflebeam), se trata de constatar la consecución de las metas fijadas por el programa y proceder a su valoración con el fin de tomar decisiones trascendentes sobre la continuidad o cambios a introducir para su mejora.

Otra de las cuestiones a resolver, previa a la selección del modelo de evaluación, consiste en la clarificación sobre cuál es el propósito de la evaluación y qué se quiere evaluar, lo que algunos autores han englobado en la evaluación de necesidades, entendiendo por estas (Tejedor 1990) las discrepancias existentes entre la situación corriente y la situación deseada, es decir el déficit, que determinará comportamientos futuros de decisión a la hora de poner en marcha un programa educativo, conforme a las prioridades señaladas por el equipo y a la asignación de recursos por las autoridades.

Asimismo es importante determinar quién debe llevar a cabo esta evaluación, Stufflebeam y Shinkfield (1989) nos hablan de un reparto de papeles en el trabajo evaluativo, así distinguen: el cliente de la evaluación (persona o grupo que utilizará los resultados), el diseñador de la evaluación (personal técnico asesor del cliente), el coordinador de la evaluación (personal especializado en tareas administrativas y de planificación), asistentes de la evaluación (personal encargado del trabajo de campo y de colaborar con otros componentes en cometidos diversos), respondientes de la evaluación (personas que rellenan formularios o proceden a la aplicación y recogida de datos), especialistas en ayuda técnica (expertos en campos de especialización -entrevistas, análisis de datos-), especialistas en información (ordenación y archivo de documentos), especialistas en comunicación (responsables de transmitir los resultados, preparar los informes), preparador de la evaluación (programas específicos de preparación para los componentes de los equipos de evaluación), investigador evaluativo (encargado de la mejora de la evaluación, investigando sobre la misma), promotor de la evaluación (función de ayuda material a los evaluadores), metaevaluador (evaluador de evaluaciones, requiere astucia y penetración).

Pretender alcanzar una composición tan ideal de equipo evaluador como la presentada resulta utópico en la situación actual de la evaluación en España, dentro del marco general de la Reforma Educativa, que si bien hace mucho hincapié en la necesidad de la evaluación a todos los niveles, después no se concreta de forma precisa en propuestas de alternativas a cada situación. No obstante, la puesta en marcha del Instituto de Calidad y Evaluación (INCE), constituye un punto de arranque sólido para dar una mayor consistencia técnica y administrativa a esta tarea, ceñida en muchas ocasiones a la mera evaluación (sanción) de los resultados académicos de los estudiantes.

---

Regresar a [Índice de materias](#)

Regresar a [Contenidos del capítulo](#)

Regresar a [comienzo de la sección](#)

---

## **II. MODELOS DE EVALUACIÓN DE PROGRAMAS:**

Dada la gran variedad de modelos existentes para la evaluación de programas, vamos a presentar de forma sintetizada los de mayor uso en el marco educativo, desde el clásico presentado por Tyler hasta la propuesta formulada por Pérez Juste. La selección de los modelos recogidos se hace teniendo en cuenta las posibilidades de aplicación al marco educativo del centro y de los proyectos diseñados por los equipos docentes de los mismos.

### **2.1. Modelo clásico de TYLER: evaluación orientada al logro:**

Los trabajos de Tyler han supuesto el punto de arranque de las propuestas de investigación evaluativa moderna. Su propuesta hace especial hincapié en que las decisiones sobre la validez o no de un programa vendrán determinadas por el nivel de congruencia entre los objetivos propuestos y los logros alcanzados, si estos eran alcanzados se tomarían las decisiones pertinentes, en cambio si no lo eran o solo se alcanzaban de forma parcial las decisiones serían distintas.

Dado el proceso lineal lógico del modelo propuesto por Tyler las fases a desarrollar en su propuesta evaluativa se concretan en las siguientes:

- Establecer metas y objetivos generales.
- Ordenar los objetivos según diversas clasificaciones.
- Definir estos objetivos en términos conductuales.
- Buscar situaciones en las que los objetivos se realizan.
- Seleccionar las técnicas apropiadas de medida.
- Recopilar los datos de la práctica.
- Comparar los datos con los objetivos conductuales fijados.

Las discrepancias que se pusieran de relieve en esta fase final deben ser corregidas y/o modificadas en los apartados correspondientes. La aportación básica del modelo consiste en desplazar la atención de la evaluación hacia otros aspectos del programa, además de los alumnos. Introduce la evaluación como referencia a criterios definidos en los objetivos. De ahí su visión recurrente de la evaluación al proporcionar información útil que permita la reformulación de los objetivos y como consecuencia la revisión de los planes evaluativos.

Las limitaciones más destacadas de esta propuesta se centran en considerar la evaluación como un proceso terminal, por lo que los posibles fallos no son detectados hasta finalizar el estudio, además el rendimiento es considerado como criterio último para decidir si se han

alcanzado o no y en qué grado las metas propuestas por el programa.

## **2.2. Modelo de STAKE: la evaluación centrada en el cliente:**

Stake recibe fuertes influencias de los modelos tylerianos, no obstante considera que estas propuestas tienen muchas limitaciones, por lo que defiende una evaluación centrada en el cliente y no solo en los resultados. Se muestra partidario del empleo de métodos cualitativos y el estudio de caso, frente al uso exclusivo de metodologías cuantitativas de base experimental. Manifiesta un gran interés por la evaluación de los denominados proyectos locales que ayuden a resolver y comprender de forma contextualizada los problemas próximos.

Su propuesta se centra en la visión de globalidad de la evaluación, por ello incluye los siguientes puntos:

1. La descripción lo más completa posible del propio programa, es decir, los antecedentes supuestos y reales, las operaciones didácticas realizadas y los resultados.
2. El juicio del programa no debe ser cometido exclusivo de los especialistas en evaluación, sino que deben participar otros colectivos (portavoces de la sociedad, profesores, padres, estudiantes), con una valoración ponderada de sus aportaciones.
3. La recopilación de datos se llevará a cabo en tres momentos, los antecedentes (información sobre el historial del grupo y sus circunstancias), las transacciones (encuentros entre los diversos colectivos implicados en el programa) y los resultados (logros de diversa índole alcanzados por el programa).

La evaluación centrada en el cliente o respondiente busca ofrecer un servicio útil y valioso a personas específicas, su característica fundamental es la sensibilidad hacia los temas esenciales, especialmente los que ayuden a los colectivos en su entorno inmediato. Por ello resalta que las evaluaciones no serán útiles si los evaluadores no conocen suficientemente tanto el lenguaje como los intereses de las personas implicadas y si los informes de evaluación no se ajustan a ese lenguaje.

Uno de los puntos básicos de la propuesta radica en la diferenciación entre la evaluación preordenada y la respondiente; así mientras la primera resalta las metas del programa como criterio de evaluación, utilizando los datos para la comprobación de los objetivos y las normas para juzgar los programas, realizando informes tipo de la investigación, es decir, se basan en un plan apriorístico; la evaluación centrada en el cliente se orienta hacia las actividades del programa, responde a necesidades de información de las audiencias y toma en consideración los diferentes valores de las personas interesadas en el programa cuando juzga su adecuación. Las propuestas y los métodos no son impuestos, sino que emergen desde la observación del programa durante el desarrollo de la evaluación.

Los principios básicos que sustentan la propuesta de evaluación de Stake son los siguientes:

- Igualdad: en la valoración de la acción de diferentes centros educativos.
- Ubicuidad: debe analizarlo todo en su globalidad.
- Diversidad: consenso explicativo dentro de las diferentes situaciones propuestas.
- Utilidad: prestar ayuda a la comunidad escolar.
- Redundancia: gran cantidad de datos sobre el programa.
- Ambigüedad: defiende la especificidad frente a la simplificación.
- Generalización: aplicabilidad de los estudios a otros casos y construcción o reformulación de teorías.

La estructura funcional operativa de la evaluación centrada en el cliente viene recogida en la obra de Stufflebeam y Shinkfield (1989, p.262). Se le denomina la evaluación del reloj por su carácter cíclico siguiendo el movimiento de las agujas y dado que contempla doce pasos:

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

Los anteriores acontecimientos o propuesta de pasos y actividades secuenciadas no tienen por qué verse necesariamente siguiendo un orden lineal; sino que algunos pueden presentarse a la vez y en ocasiones es preciso volver sobre otros. La propuesta estratégica general formulada por Stake hemos de buscarla en términos de utilidad y legitimidad, por ello opina que sirve tanto a la evaluación formativa como a la sumativa. A la vez que reconoce que las valoraciones de la evaluación preordenada son más objetivas y exactas. No obstante, por encima de otras consideraciones, estima que toda evaluación debe ser adaptable y la evaluación respondiente es claramente superior en el cumplimiento de este principio.

### **2.3. Modelo de STUFFLEBEAM: evaluación orientada a la decisión:**

El modelo se sustenta sobre la base de la importancia de la toma de decisiones sobre un determinado programa con el fin de maximizar la utilización de los resultados de la evaluación, no solo al finalizar el proceso, sino a lo largo del mismo. De la definición de evaluación apuntada anteriormente se extraen tres objetivos básicos: servir de guía para la toma de decisiones, proporcionar datos y promover la comprensión de los fenómenos estudiados. La evaluación se presenta como un proceso en el que se incluyen las etapas de identificar, obtener y proporcionar información. De ahí su propuesta CIPP, que distingue cuatro tipos de evaluaciones con sus correspondientes decisiones de mejora:

- Evaluación del contexto: sirve de base a las decisiones de planificación, necesidades de un programa para definir sus objetivos.
- Evaluación del "input" o entrada: nos indica los recursos disponibles, los métodos aplicables, las estrategias alternativas al programa y los planes de mayor potencial.
- Evaluación del proceso: proporciona información sobre la marcha del programa y si las actividades siguen el plan previsto, con el fin de introducir las modificaciones que sean necesarias.
- Evaluación del producto: valorar, interpretar y juzgar los logros del programa. Analizar si se han alcanzado los objetivos propuestos para decidir si el programa debe seguir o ampliarse o reformularse.

Uno de los puntos básicos de referencia del modelo CIPP nos viene determinado por su utilización en el perfeccionamiento de los sistemas, tal y como se recoge en la propuesta del cuadro tomado de (Stufflebeam y Shinkfield 1989, p. 192)

1. Operaciones regulares del sistema (incluyen distintas actividades evaluativas).
2. Evaluaciones periódicas del contexto, de ello se derivará la justificación de un cambio, si la opción es negativa retomamos el punto 1. Si la opción es positiva definimos el problema y formulamos los objetivos. Si existe una solución satisfactoria aplíquese y vuelva al punto de partida. Si no la hay pasamos al punto 3.

3. Evaluación de entrada, nos valora las estrategias encontradas, si no son adecuadas retomamos el punto 2. En cambio, si son prometedoras debemos avanzar en su desarrollo y comprobación, mediante la propuesta de realización de un proyecto especial.

4. Evaluación del proceso y del producto, se centrará en valorar los rendimientos del sistema, si éstos son satisfactorios se debe aplicar la solución propuesta. En caso contrario, debemos plantearnos si el proyecto merece un nuevo intento, lo que nos lleva a retomar la parte final del punto 3; o bien, al fracaso de la operación inicial planteada.

Del análisis de dicha propuesta y circunscrita la explicación a un contexto de reforma o cambio institucional nos ofrece la siguiente aportación, siguiendo las directrices emanadas de la formulación del modelo:

a. La evaluación debe formar parte de todo programa regular de cualquier institución educativa y no circunscribirse en exclusiva a una actividad especializada relacionada con proyectos de innovación o renovación.

b. La evaluación desempeña un papel de suma importancia tanto en lo que se refiere a la estimulación de los sujetos implicados en la misma como en lo relativo a la planificación de cambios.

c. El empleo de cada uno de los cuatro tipos de evaluación propuestos en el modelo CIPP es necesario cuando nos permitan obtener una mayor información, es decir actúan como partes del gran marco evaluativo de cada institución, siendo su función básica el analizar aspectos marginales de suma importancia para el proyecto.

d. La implementación de nuevos programas en el marco educativo debe prever el empleo de la evaluación, una vez que hayan sido aceptados por la comunidad.

e. La información suministrada además de resolver problemas de tipo institucional proporciona una base científica para juzgar si las decisiones sobre el fracaso o el éxito de un programa especial se toman desde posiciones debidamente justificables.

f. Las decisiones del equipo evaluador acerca del principio, mantenimiento, aplicación o final de los programas educativos o relativas a su perfeccionamiento son el reflejo de las fuerzas dinámicas que van más allá de la esfera próxima de estudio e influencia directa del evaluador.

En consecuencia podemos afirmar que el modelo CIPP nos proporciona diferentes tipos de decisiones que nos ayudan a crear las estrategias evaluativas apropiadas, entre ellas destacan:

- Decisiones de realización: sobre la puesta en marcha del programa (recursos, apoyos humanos).

- Decisiones de reciclaje: continuar o no con el programa, unirlo a otro, buscar recursos nuevos.
- Decisiones de planificación: selección de metas valiosas para el programa y los sujetos implicados.
- Decisiones de estructura: medios necesarios para alcanzar esas metas.

Sin duda la propuesta de Stufflebeam en una de las más clarificadoras y ambiciosas en la evaluación de programas educativos, el mismo autor reconoce algunas de las limitaciones para llevar a cabo una evaluación exhaustiva siguiendo su modelo. Pero sí recomienda encarecidamente que la evaluación no es un fin en sí misma, sino que se incardina dentro de un proceso educativo más amplio y globalizado, por ello propugna que la propia evaluación sea objeto de evaluación (metaevaluación) recurriendo a una serie de criterios, procesos y técnicas apropiados para su cometido.

#### **2.4. Modelo de SCRIVEN: la evaluación orientada hacia el consumidor:**

El modelo propuesto por Scriven parte de una crítica a otros enfoques excesivamente centrados en alcanzar los objetivos de los participantes, mientras que según este autor se deben satisfacer prioritariamente las necesidades de los consumidores. Define la evaluación como "la determinación sistemática y objetiva del valor o el mérito de algún objeto". Por ello defiende la objetividad de los criterios de valoración, el desarrollo de la lógica probativa, para culminar el proceso en la emisión de juicios y recomendaciones a los consumidores. Asimismo concede una gran importancia a las necesidades sociales para la ciencia del valor postulada.

Para Scriven la evaluación cumple dos funciones básicas: la formativa (ayuda a desarrollar programas) y la sumativa (determina el valor del objeto una vez que se ha desarrollado); que van a dar origen a los conceptos evaluación formativa y evaluación sumativa empleados por el autor.

La *evaluación formativa* centrará su cometido en la determinación de las actividades y resultados sobre la marcha o durante el proceso, a la vez que nos indica hasta qué punto están funcionando o han funcionado bien los procesos y los procedimientos para alcanzar o satisfacer determinados objetivos. En general proporciona ayuda al personal para alcanzar el perfeccionamiento.

La *evaluación sumativa* se puede llevar a cabo al finalizar una unidad, un semestre, un curso o un ciclo. En ella se analizan todos los efectos de los objetos y se comparan con las necesidades de los consumidores para tomar las decisiones oportunas. Suele ser realizada por un evaluador externo, además se toman como base las aportaciones de la evaluación formativa. En síntesis proporciona valoraciones independientes que nos permiten comparar los costes, los méritos y los valores de los distintos programas.

Algunas de las aportaciones más significativas apuntadas por Scriven hacen alusión a la distinción entre la evaluación amateur (autoevaluación) en las primeras fases, reservando la evaluación profesional para las últimas etapas del desarrollo del programa. Asimismo establece una distinción entre la evaluación intrínseca (cualidades de la mediación) y la evaluación final (efectos sobre los clientes). Prefiere una propuesta de evaluación sin metas, lo que lleva al evaluador a analizar todos los efectos del programa y no solo aquellos que traten de ratificar o no la consecución de unos objetivos, este enfoque se considera menos intrusivo, más adaptable a los cambios, fomenta el hallazgo de efectos secundarios, mas estimulante profesionalmente y más equitativa en la visión social de los valores. Además se concede una gran importancia a la valoración de las necesidades y las repercusiones que ello puede tener sobre el desarrollo del programa, lo que le permitirá juzgar los resultados del mismo como buenos, malos o indiferentes, ello implica una gran complejidad práctica.

La lista de control de indicadores de Scriven que ha tenido diversas actualizaciones en los últimos años se concreta en los puntos siguientes:

1. Descripción: qué es lo que se va a evaluar.
2. Clientes: quién encarga la evaluación.
3. Antecedentes y contexto: del objeto evaluado y de la evaluación en sí misma.
4. Recursos: disponibles para su uso por los colectivos implicados.
5. Funciones: que hace el evaluando.
6. Sistema de distribución: qué se consigue y cómo se alcanza.
7. El consumidor: quién está utilizando o recibiendo los efectos.
8. Las necesidades y los valores: de la población afectada real o potencialmente.
9. Normas: previamente validadas, de mérito o valor a aplicar.
10. Procesos: limitaciones de costes, beneficios de las operaciones normales.
11. Resultados: efectos producidos.
12. Generalización: a otras personas, lugares, tiempos o versiones.
13. Costes: económicos, psicológicos, personales. Inicial, repeticiones. Directos, indirectos.
14. Comparaciones: con otras opciones alternativas.
15. Significado: síntesis de todo lo anterior.

16. Recomendaciones: sean o no requeridas y puedan llevarse a cabo o no en situaciones reales.

17. El informe: cuidar el vocabulario, la extensión, el diseño, tiempo, localización, personal, presentación.

18. La meta-evaluación: debe realizarse antes de la realización y la difusión del informe.

Estas propuestas no necesariamente deben llevarse a cabo de una forma secuencial, si bien, una vez recogidos los datos, analizados y sintetizados, el evaluador puede introducir sus tendencias en el informe final y someterlas a crítica antes de la redacción definitiva del informe de evaluación.

## **2.5. Modelo de PÉREZ JUSTE: evaluación integral e integrada:**

Partiendo del concepto de evaluación de programas apuntado al comienzo de la exposición cabe resaltar como notas básicas de esa definición la idea de proceso incardinado en algo más amplio que lo meramente medicional, además sistemático lo que implica una planificación, diseñado de forma intencional y técnica, esto es, reconociendo explícitamente el papel del evaluador y las exigencias técnicas de esta tarea. La recogida de información supone una superación de los modelos restrictivos de prueba y ampliar los datos a efectos no planeados inicialmente, además esta información deber ser rigurosa (exige planteamientos científicos), válida (riqueza de datos) y fiable (precisa en sus aportaciones). El concepto incluye el juicio de valor lo que nos lleva a explicitar tanto los criterios como las referencias para analizar la calidad y los logros del propio programa educativo. Todo ello dirigido al objetivo último de la toma de decisiones que repercutan sobre la mejora no solo centradas en el programa educativo sino en el personal implicado en su realización, lo que repercutirá de forma positiva sobre el grupo social sobre el que se está trabajando.

La propuesta del modelo evaluativo recoge tres momentos que a su vez comprenden unas dimensiones con unos objetos concretos que precisan de los correspondientes criterios e indicadores de evaluación. En dicha propuesta se pone de relieve la importancia de la evaluación integral de todos los elementos que intervienen en el proceso educativo e integrada en un marco de referencia más amplio como es el centro escolar o el propio sistema educativo.

De forma esquemática los puntos propuestos por Pérez Juste (1994) en su modelo son los siguientes:

**1. Momento inicial (el programa en sí mismo):** que comprende tres dimensiones:

- Calidad intrínseca del programa:

\* Contenido del programa.

- \* Calidad técnica.

- \* Evaluabilidad.

- Adecuación al contexto:

- \* Respuesta a necesidades y carencias.

- \* Priorización.

- Adecuación a la situación de partida:

- \* Viabilidad

**2. Momento procesual (el programa en su desarrollo):** comprende dos dimensiones:

- Ejecución:

- \* Actividades.

- \* Secuencias.

- \* Tiempo.

- \* Flexibilidad.

- Marco:

- \* Clima.

- \* Coherencia.

**3. Momento final (el programa en sus resultados):** comprende tres dimensiones:

- Medida y logros:

- \* Constatación.

- \* Contraste.

- Valoración:

- \* Criterios.

- \* Referencias.

- Continuidad:

- \* Decisiones.

\* Incorporación de las mejoras.

\* Plan de seguimiento.

Esta propuesta es muy ambiciosa en sus planteamientos, si bien como el mismo autor reconoce, no se trata de llevarla a cabo en un contexto determinado de forma exhaustiva, sino que, puede ofrecer unas pautas rigurosas de actuación a los evaluadores en el marco escolar. Quizá una de las aportaciones principales a destacar radica en el rigor con que se trata la evaluación inicial (el programa en sí mismo), ello permitirá conocer, antes de someter a aplicación, las virtualidades y los posibles defectos del programa, de tal forma que se puedan tomar decisiones de mejora del mismo. Por ello, destacaríamos el marcado matiz pedagógico del modelo, traducido a lo largo de toda la propuesta y de forma más explícita en la formulación de los criterios e indicadores de la evaluación formulados para cada uno de los momentos y dimensiones.

## **2.6. Consideraciones finales:**

Las anteriores propuestas de evaluación de programas educativos abarcan un amplio espectro informativo para que el docente, en el contexto específico de su programa, aula o centro, pueda llevar a cabo una evaluación sólida y científica. La opción elegida no necesariamente debe coincidir en sus extremos con alguno de los modelos, al contrario, las diferentes aportaciones pueden enriquecer el trabajo y ofrecernos una visión más crítica y objetiva de la realidad educativa analizada.

No obstante, la propia realidad del profesorado en su centro educativo, con una sobrecarga en la actividad docente directa en el aula ante sus alumnos, hace que el resto de posibles tareas educativas se vean desplazadas en aras a cubrir unas necesidades prioritarias.

Desde esta visión objetiva, hacemos la propuesta, para que sirva como reflexión a los docentes en ejercicio, sobre los diferentes cometidos del área evaluativa, que completen la clásica evaluación de los aprendizajes de los estudiantes, cometido casi exclusivo del profesorado en estos niveles educativos.

---

Regresar a [Índice de materias](#)

Regresar a [Contenidos del capítulo](#)

Regresar a [comienzo de la sección](#)

---

### **III. TÉCNICAS E INSTRUMENTOS PARA LA EVALUACIÓN DE PROGRAMAS:**

Los diferentes modelos comentados con anterioridad sobre la evaluación de programas educativos hacen especial hincapié en la trascendencia de alcanzar unos buenos datos y de calidad, de forma que nos permitan garantizar, con un alto grado de acierto, la idoneidad de las etapas planteadas en dichas propuestas. Para ello, es preciso que en cada una de las fases o momentos de la evaluación de programas seamos conscientes de la importancia de las mismas en su desarrollo posterior.

El estudio sobre la realidad educativa puede ser abordado desde un posicionamiento racionalista o cuantitativo, o bien desde una óptica naturalista o cualitativa, lo que da origen a los grandes paradigmas o modelos de análisis. La elección de uno u otro va a repercutir de forma directa sobre los procedimientos de estudio que guíen la investigación evaluativa.

Así pues, mientras el método es el camino a seguir, las técnicas nos indican el modo de recorrer ese camino, los procedimientos que deben seguirse. Por ello, prima su carácter práctico y operativo de proximidad a la realidad a evaluar. En este sentido las técnicas serían aquellos medios que emplea el investigador para registrar sus observaciones o facilitar el tratamiento experimental.

Por su parte los instrumentos configuran los aparatos que nos facilitan la recogida de datos en la evaluación de programas educativos. Las clasificaciones aportadas por diversos autores (Pérez Juste, Bisquerra, Ander-Egg, Van Dalen), no coinciden en una diferenciación clara entre instrumentos y técnicas, de ahí que lo que unos consideran como técnica para otros autores se trata de un instrumento.

Dada la complejidad metodológica del problema, nuestra propuesta de clasificación para las técnicas e instrumentos de evaluación se concreta en los siguientes puntos:

#### a. Instrumentos de medición:

- Tests:
  - \* Aptitudes
  - \* Rendimiento
  - \* Personalidad
- Escalas.

#### b. Métodos de observación:

- No estructurados:

- \* Observación participante
- \* Registro de especímenes
- \* Anecdóticos
  
- Estructurados (listas de comprobación).

c. Instrumentos para el registro de datos:

- Ficha anecdótica
  
- Listas de control
  
- Escalas de actitudes y valores
  
- Escalas de estimación:
  - \* Numéricas
  - \* Gráficas
  - \* Verbales o descriptivas.

d. Cuestionarios y entrevistas:

- Cuestionarios:
  - \* Individuales o grupales
  - \* Aplicación personal o envío por correo
  - \* Cerrados o abiertos
  
- Entrevistas:
  - \* Estructurada o formal
  - \* No estructurada o informal: Focalizada, Clínica, No dirigida.

e. Análisis de tareas y documentos:

- Tareas escolares:
  - \* Tabla de criterios
  - \* Adjudicación de valores

- Escalas de producción escolar:

- \* Escritura
- \* Ortografía
- \* Dibujo
- \* Composición

- Análisis de contenido.

f. Métodos indirectos:

- Proyectivos:

- \* Gráficas
- \* Verbales
- \* Juego
- \* Psicodrama y socio drama

- Más estructurados:

- \* Tests de conocimientos y habilidades
- \* Tests de información
- \* Juicios de actitud relevante.

g. Otras técnicas:

- Metodología "Q"
- Diferencial semántico
- Técnicas socio métricas.

Desde la perspectiva concreta de la evaluación de programas educativos y partiendo del modelo formulado por Pérez Juste, vamos a presentar alguna de las técnicas más empleadas en las primeras fases de la evaluación, para recoger también algún instrumento especialmente útil en la última fase del proceso.

### **3.1. Técnicas para la evaluación de programas:**

En este punto vamos a presentar tres de las principales técnicas empleadas en la evaluación de programas educativos, además en cada una de ellas trataremos de analizar su contribución específica en cada una de las fases del proceso evaluador.

### **3.1.1. La observación:**

El observador contempla la realidad sin actuar sobre ella, registra hechos o fenómenos tal y como se están produciendo en un momento determinado. Ahora bien, debemos distinguir lo que puede ser una observación casual de los fenómenos naturales que ocurren a nuestro alrededor de aquella otra perfectamente planificada y sistemática, que alcanza el rango de cientificidad.

SELLTIZ (1980, p.346) nos dice que la observación se convierte en una técnica científica cuando cumple los siguientes requisitos:

- Sirve a un objeto de investigación previamente formulado.
- Se planea de forma sistemática.
- Se registra sistemáticamente y se relaciona con proposiciones más generales.
- Se somete a comprobaciones y controles sobre su validez y fiabilidad.

La observación es aplicable a cualquier hecho o fenómeno educativo, pero si se realiza de forma indiscriminada perdería la mayor parte de su interés, se hace preciso seleccionar el objeto o el tema a observar. Así DE KETELE (1984, p.21) la define como "un proceso que requiere atención voluntaria e inteligencia, orientado por un objetivo terminal y organizador, dirigido hacia un objeto con el fin de obtener información".

Así pues, la intencionalidad es una de sus características básicas, ello implica determinar con claridad lo que se va a observar, como se va a llevar a cabo y quién es el encargado de la misma. En primer lugar debemos tener perfectamente delimitado el campo de la observación, seleccionando puntos concretos, pues tratar de observar la realidad total es prácticamente imposible, ello se traduce operativamente en lo que llamamos guión de la observación. En segundo lugar se nos plantea el cómo observar, bien con la presencia directa del observador o bien recurriendo a diversos recursos audiovisuales y de registro de la actividad, en cualquier caso debemos cuidar los problemas de interacción entre el observador y el observado. En tercer lugar tendríamos que definir quién debe llevar a cabo la misma, la persona que refleja en el correspondiente registro y toma las anotaciones pertinentes, puede ser una persona ajena (externa) o bien alguien dentro del grupo (interna).

Unido a la intencionalidad cabe resaltar la estructuración o la vinculación a unos determinados presupuestos teóricos y realizada con orden y método apropiado. También debemos señalar la importancia del control, de forma que se eliminen en el mayor grado

posible la influencia de factores externos que puedan distorsionar la visión de la realidad observada.

Existen diferentes clasificaciones de la observación, según se contemplen los medios empleados, el número de observadores, el lugar en que se realiza. Desde la perspectiva de la evaluación de programas nos parece más interesante aquella que hace referencia al grado de participación del observador. Desde este enfoque vamos a distinguir dos modalidades:

- La observación externa o no participante.

- La observación interna o participante.

a) **La observación externa o no participante** se caracteriza por la no pertenencia del observador al grupo que se estudia, procura contemplar la realidad desde fuera. Esta tarea se puede llevar a cabo de forma *directa*, sobre el terreno y en contacto inmediato con la realidad; o bien, de forma *indirecta* basándose en datos estadísticos (censos) y fuentes de tipo documental (prensa, archivos...). Como ventaja de esta modalidad cabe reseñar la especial dedicación del observador a la tarea, sin interferencias de otro tipo. Como inconveniente señalaremos el que solo nos permite percibir los aspectos más externos de la vida social, además de que exige el consentimiento de la persona o grupo que son objeto de la observación y que puede tener repercusiones negativas sobre la validez de los datos.

b) **La observación interna o participante**, también denominada observación activa consiste en la participación directa del observador que asume diferentes roles en la vida del grupo, en su organización, conversando con todos los componentes, de forma que su presencia no perturbe o interfiera el desarrollo natural de los acontecimientos. Es decir, se llega a conocer la vida de un grupo desde el interior del mismo. Ello significa que el observador además de ver lo que está aconteciendo, lo siente y se integra en la situación o en el desarrollo del programa. Debe combinar los roles de observador y participante, lo que le permitirá comprender el programa desde un punto de vista interno, a la vez que lo describe para los que están fuera de él. Si el observador pertenece al grupo que se investiga nos encontramos ante una participación natural; ahora bien, si el observador solamente se integra en el grupo con el objetivo de llevar a cabo una investigación hablaremos de participación artificial.

### **3.1.2. La entrevista:**

Es uno de los procedimientos más empleados en la investigación social y educativa, dado que nos permite una recogida directa de datos sobre la base de una relación personal entre entrevistador y entrevistado, mediante anotaciones y comentarios no solo sobre lo que se busca directamente, sino también sobre las circunstancias en que estos hechos se producen, que pueden ofrecernos una información rica y valiosa en la interpretación de los datos recogidos.

Por tanto, la entrevista cumple básicamente las siguientes funciones: obtener información de los sujetos o los grupos, facilitar información a otras personas e influir sobre

determinados aspectos de la conducta. En realidad abarca las funciones: diagnóstica, investigadora, orientadora y terapéutica.

Desde la perspectiva de la evaluación de programas, cumple un importante cometido en la recogida de información. Así FOX la sitúa entre los recursos que nos permiten recoger datos superficiales, subsuperficiales y profundos. Ya sean sobre cuestiones externas al sujeto que requieren su opinión, como sobre aspectos internos identificados con las experiencias, propósitos e ideales del sujeto entrevistado. Por su parte KERLINGER (1981, p.338) le atribuye tres finalidades básicas: identificar variables o sugerir hipótesis, ser el instrumento de recogida de datos de una investigación evaluativa y validar otros métodos o permitir la profundización en las motivaciones y respuestas de los entrevistados.

Para garantizar la fiabilidad y la validez de los datos recogidos en la entrevista es conveniente que tengamos en cuenta una serie de aspectos claves en esta técnica, entre otros vamos a señalar los siguientes:

- La entrevista supone una relación humana, por tanto, debemos conseguir una comunicación abierta, espontánea, positiva, libre de tensiones, dentro del marco de la flexibilidad propia de esta técnica. Ello exige alcanzar un alto nivel de "empatía" (simpatía neutral) que se manifiesta en la comprensión de los demás, la intuición de la vida ajena, pero también en no juzgar ni implicarse afectivamente tratando de alcanzar los niveles máximos de objetividad.
- El contenido de la entrevista, pues no todos los sujetos tienen la suficiente preparación para contestar determinados temas complejos, o bien dado su nivel de implicación personal resulta difícil conseguir la objetividad de las respuestas. En aquellos casos en que busquemos la opinión de un determinado colectivo deberemos optar por una muestra representativa de la población objeto de estudio. En otras ocasiones debemos recurrir, dado lo delicado del tema de estudio (creencias, sentimientos, actitudes), a preguntas indirectas.
- Las unidades de información que configuran la entrevista, en torno a las cuáles se agrupan las preguntas y respuestas, de las que va a depender en gran medida los niveles de fiabilidad y validez de la información recogida. Por ello es preciso determinar el tipo de preguntas y respuestas más apropiado (cerradas, abiertas y semiabiertas), la formulación clara y concisa de las cuestiones (si es preciso acompañando las explicaciones pertinentes de temas excesivamente técnicos), su estructuración uniforme o no para todos los sujetos. Sería conveniente garantizar la libertad del entrevistador.
- El registro de las respuestas debe garantizar la fidelidad de las expresiones manifestadas por los sujetos, evitando posibles alteraciones interpretativas. Por ello en determinadas ocasiones además del lápiz y papel hemos de recurrir a los medios audiovisuales como las cassettes o los videos, siempre con el consentimiento de la persona entrevistada. Ofrece mayores garantías de objetividad de los datos y en caso de duda se pueden volver a consultar; como contraprestación los comportamientos naturales se suelen alterar con la presencia de estos medios.

- Finalmente es preciso proceder a la categorización y valoración de la información recibida, por ello debemos identificar todas las categorías de la realidad estudiada, que éstas sean mutuamente excluyentes evitando solapamientos e interferencias. Traducido a valores numéricos cuantitativos nos permitirán llevar a cabo los preceptivos análisis estadísticos y comprender la realidad estudiada.

Aunque existen diferentes criterios en la clasificación de la entrevista, si tenemos en cuenta el grado de estructuración o la directividad, estableceremos la siguiente distinción:

- Entrevista tipificada o estructurada

- Entrevista no estructurada o flexible:

\* Entrevista focalizada

\* Entrevista clínica

\* Entrevista no dirigida.

a) **Entrevista estructurada**, es aquella que emplea un protocolo previo de preguntas de alternativa fija o bien de varias alternativas de respuesta para que el sujeto elija la mas apropiada. En realidad se trata de un interrogatorio en el que las preguntas, previamente establecidas, se formulan en el mismo orden y con los mismos términos. Tanto las preguntas como las respuestas se reflejan de forma escrita o mediante la correspondiente codificación.

b) **Entrevista no estructurada**, permite una mayor libertad a la iniciativa tanto del entrevistador como del entrevistado, la conversación permite y estimula una expresión más libre y con respuestas no determinadas de antemano, dado que las preguntas que se formulan suelen ser de tipo abierto. Dentro de esta modalidad debemos distinguir tres opciones diferentes en cuanto a los campos concretos de la entrevista.

b.1. *Entrevista focalizada*, se trata de una modalidad muy técnica, que exige del entrevistador una gran habilidad, experiencia y tacto personal. Parte de una lista de tópicos elaborada por el entrevistador, en torno a los cuáles se focaliza la entrevista. Si bien, éste dispone de libertad en el proceso de aplicación. Dado su carácter precisa de agudeza y habilidad por parte del entrevistador para buscar aquello que realmente desea, centrando el interrogatorio sobre las cuestiones más precisas, sabiendo escuchar al entrevistado, ayudándole a expresarse de forma correcta y, si ello fuera preciso, esclareciendo los términos confusos, pero sin llegar a sugerir las respuestas.

b.2. *Entrevista clínica*, se diferencia de la anterior en el enfoque, dado que se centra en el análisis de las motivaciones y sentimientos de los sujetos. En el guión de la entrevista se deben reflejar las cuestiones básicas que deben ser abordadas en el desarrollo de la misma.

b.3. *Entrevista no dirigida*, en este caso el entrevistador, tomando como referencia el problema objeto de estudio, se acerca a su estudio actuando como un catalizador, facilitando el lenguaje básico del entrevistado. Este tiene completa libertad para expresar

sus sentimientos y opiniones, mientras que el entrevistador debe ser capaz de crear una atmósfera facilitadora para que el sujeto pueda expresarse con total libertad. Ello permite recoger datos tanto sobre hechos de comportamiento externo y observable como sobre motivos e ideales relacionados con la intimidad de la persona.

La entrevista configura un medio rico y valioso de recogida de información, puesto que no solo se recoge ésta, sino que se completa con otras circunstancias que sirven de contexto y marco de referencia de aquella, lo que enriquece y posibilita su interpretación. Por ello, es preciso cuidar los aspectos técnicos de la misma como pueden ser la **relevancia de los datos**, que nos debe garantizar que las preguntas o ítems que hemos preparado sean los más apropiados al problema de estudio. Respecto a la **fiabilidad** de las respuestas es preciso tomar precauciones según la información buscada, así cuando se trate de pensamientos, ideas, sentimientos, la fiabilidad se manifiesta en el grado de congruencia entre las distintas respuestas, en cambio, cuando se trate de datos sobre un mismo hecho recogidos de diferentes informadores debemos analizar la concordancia entre los mismos y la consistencia interna de la información recogida. En cuanto a la **validez**, está muy ligada a la relevancia de la información y también a la fiabilidad. Exige que las respuestas guarden relación con el tema estudiado, además de representar la información deseada y no otra (validez de contenido), para ello debemos recurrir a la consulta a jueces o expertos.

### **3.1.3. El análisis de contenido:**

Es una técnica que nos ofrece la oportunidad de investigar sobre las producciones personales, sean o no elaboradas intencionalmente para una investigación concreta. Se trata de analizar y cuantificar los diferentes materiales de comunicación ya sean éstos escritos (prensa, libros, revistas, documentos) o audiovisuales (radio, cine, televisión).

BERELSON (1952, p.18), uno de los pioneros en el uso de esta técnica la define como "técnica de investigación para la descripción objetiva, sistemática y cualitativa del contenido manifiesto de las comunicaciones con el fin de interpretarlas". Por su parte KRIPPENDORF (1990, p.28) lo conceptúa como "técnica de investigación destinada a formular, a partir de ciertos datos, inferencias reproducibles y válidas que puedan aplicarse a un contexto".

El análisis de contenido admite dos niveles de análisis: el manifiesto o lo que aparece a simple vista y el latente o lo que subyace o puede leerse entre líneas. En sus comienzos la técnica se limitaba al estudio del primer nivel, se analizaba lo que estaba presente, es decir los contenidos objetivos de la información. Con posterioridad se consideraron pobres estos análisis y se pasó al segundo nivel sobre las intenciones o motivaciones de quién facilita la información, se analizan las implicaciones y derivaciones de las respuestas. La mayor parte de los autores se inclinan por el análisis de lo manifiesto, que servirá de guía para abordar el estudio de lo latente.

En la realización del análisis de contenido, como técnica eminentemente descriptiva, se sigue el siguiente proceso:

a) *Precisar el objetivo que se persigue*, para poder orientar de forma adecuada el trabajo en sus orígenes. La definición clara y concisa de estos objetivos va a encaminar el resto del proceso de investigación de la técnica.

b) *Definir la población o universo de estudio*, tanto en lo que se refiere a la unidad de contenido a estudiar como en decidir el cuánto de dicho análisis, se trata de establecer una limitación temporal o selectiva de documentos (muestreo).

c) *Establecer las unidades de análisis*, que constituyen los núcleos con significado propio que serán objeto de estudio para su clasificación y/o recuento. Sin duda configura el punto clave para que el análisis de contenido ofrezca una información válida y fiable al investigador. Estas unidades pueden ser: la palabra, la frase, la oración o el párrafo. En otras ocasiones la unidad es el tema presentado, también se suele utilizar el espacio ocupado (página, columna, lugar).

d) *Determinar las categorías del análisis*, se trata de la decisión más difícil y complicada en el empleo de esta técnica, pero a la vez la más decisiva y creativa. La categorización consiste en la clasificación de los elementos de un conjunto a partir de criterios lógicos, psicológicos y contextuales previamente definidos. En todo caso debemos tener presentes las reglas lógicas de: homogeneidad, utilidad, exclusión mutua y claridad y concreción. Por ejemplo contenidos y formas de abordarlos: hechos, comentarios, pruebas científicas, opiniones, descripciones, valoraciones, intensidad, cantidad, calidad...

e) *Cuantificar las diferentes categorías*, para facilitar el posterior análisis de los datos. Una de las más empleadas consiste en determinar la frecuencia del tipo de unidad de análisis elegido. También se suelen emplear los rangos o posiciones de orden de los datos. En consecuencia los estadísticos prioritarios de uso son los no paramétricos, si bien, las técnicas estadísticas multivariantes abren un amplio abanico de posibilidades de abordar análisis más complejos mediante el uso del ordenador.

f) *Interpretar los datos y elaborar conclusiones*, constituyen la fase final de todo proceso investigador. Así la interpretación de los datos requiere contrastar los valores empíricos con los valores críticos de distribución que nos permitan establecer la existencia o no de diferencias y su significatividad estadística. Sobre esta base debemos elaborar nuestras conclusiones y aportaciones de la técnica a la resolución de los objetivos propuestos inicialmente y si éstos se han alcanzado de forma satisfactoria.

Para concluir haremos una breve alusión a los aspectos técnicos del análisis de contenido, así en lo que se refiere a la fiabilidad esta la debemos entender como la estabilidad de los resultados, la reproductibilidad y la exactitud. Mientras que la validez, en su sentido interno, se basa en la existencia de una relación lógica entre indicadores, categorías y respuestas. Completándose con la validez de contenido, mediante la consulta a jueces o expertos en el tema estudiado.

### **3.2. Instrumentos de recogida de datos en la evaluación de programas:**

Dada la amplitud y complejidad de los instrumentos de recogida de datos y de medida empleados en la evaluación de programas educativos y las limitaciones de la propuesta, vamos a presentar algunos de los de mayor uso en el campo que nos ocupa y de fácil construcción y uso en el marco escolar.

### **3.2.1. El cuestionario:**

Consiste en un conjunto de preguntas o cuestiones de uno o varios tipos que se consideran relevantes para el rasgo, la característica o la dimensión sobre la que se desea obtener información, puede ser aplicado de forma variada: individual, en grupo, presencialmente, por correo. En la mayor parte de los casos se da una escasa relación entre los sujetos que contestan y la persona que los administra. Debe suponer un intento de recoger toda la información necesaria y relevante para el problema investigado, pero solo esa información. Es necesario garantizar que los sujetos no falsean las respuestas para quedar bien ante el administrador, por ello, debemos respetar la intimidad de la persona que contesta.

Construir un buen cuestionario requiere un proceso difícil y laborioso, TENBRINK (1981, p.275) apunta los siguientes pasos: describir la información que se necesita, redactar las preguntas, ordenar las preguntas, ofrecer un medio para responder, escribir las instrucciones y reproducir el cuestionario. De forma más sintética los aspectos básicos a considerar son los siguientes:

a) *El contenido* de los cuestionarios se dirige a dos grandes campos, uno configurado por las cuestiones relativas a la identificación de los sujetos desde diferentes variables relevantes para el estudio y otro constituido por las opiniones, hechos, actitudes de los sujetos ante diferentes propuestas del investigador. En este apartado debemos combinar la relevancia y la pertinencia de la información con la longitud del cuestionario (posible fatiga, riesgo de abandono o falta de seriedad en las respuestas).

b) *Las preguntas y respuestas*, están íntimamente enlazadas en la elaboración del cuestionario, pues la formulación de las primeras condiciona las respuestas de los sujetos. La pregunta debe lograr la colaboración del encuestado, por ello debe quedar claramente especificada la intención del investigador, además de la exigencia de un lenguaje claro y adecuado a las características de los sujetos, con enunciados concretos y precisos que eviten dudas o interpretaciones erróneas, procurando buscar la coherencia en su estructura gramatical y las formulaciones progresivas de las mismas, huyendo de respuestas sugeridas o tendenciosas. Asimismo debemos cuidar que las diferentes alternativas formuladas incluyan todas las posibilidades de respuesta.

Referente a las respuestas pueden ser totalmente abiertas o libres en las que es el sujeto el que decide qué, cómo, cuánto y a qué contestará, elaborando personalmente su propia respuesta. Una segunda opción sería ofrecer un abanico de posibles respuestas para que sean valoradas y añadir un final abierto dónde se puede completar o ampliar la información solicitada. La tercera posibilidad sería la totalmente estructurada en la que el sujeto se

limita elegir la mejor o mejores respuestas o bien a valorar determinados items mediante una escala.

c) *La forma de recoger la información*, debe garantizar la fidelidad y precisión de los datos recogidos ya sea de forma personal o mediante recursos tecnológicos (magnetófonos, videos), en las aplicaciones masivas o envíos por correo el sujeto refleja sus respuestas en el mismo cuestionario o en la hoja de aplicación especialmente diseñada para este fin.

Al margen de otras consideraciones de tipo práctico referentes a la aplicación misma del cuestionario uno de los temas más complicados consiste en la determinación de las características técnicas de este instrumento. La determinación de la fiabilidad resulta muy compleja, por ello en ocasiones se nos habla de indicadores de fiabilidad evaluados mediante la sinceridad y concordancia de las respuestas ofrecidas a items similares o estrechamente relacionados. No obstante, nos deben preocupar más los aspectos relacionados con la validez, en el sentido de garantizar que las preguntas recogidas en el cuestionario logran una presencia suficiente y representativa de los objetivos y contenidos de la investigación que estamos planteando, además de valorar si su ubicación es la correcta, para ello debemos recurrir a la opinión de los expertos o bien mediante una aplicación piloto determinar estas características.

### **3.2.2. Las escalas:**

Se trata de instrumentos para evaluar diversos aspectos de los fenómenos sociales, en la forma más objetiva posible. KERLINGER (1981, P.345) define la escala como "una serie de símbolos o números contruidos de tal modo que puedan ser asignados según determinadas reglas a los individuos (a sus comportamientos) a quienes se aplica". Aunque existen diversos tipos de escalas, nosotros vamos a distinguir dos modalidades: las de estimación y las de actitudes.

a) *Escalas de estimación* recogen las apreciaciones del sujeto sobre el grado en que se manifiesta un comportamiento. Es decir se asigna un valor numérico al punto o categoría en que el clasificador sitúa a la persona u objeto analizado.

a.1. Escala gráfica: se coloca una señal en el punto elegido que va de un extremo a otro del atributo propuesto.

Blanco ..... Negro

a.2. Escala numérica: sobre la misma base pero con números ( 1 a 5, 1 a 7, 1 a 9).

a.3. Escalas descriptivas o de categorías específicas: exigen una definición específica de cada una de ellas en una escala de tres, cinco o siete, para que el sujeto elija aquella más acorde con su opinión o situación.

En algunas ocasiones la escala construida por el investigador tiende a combinar los modelos anteriores, de manera especial la tercera con cualquiera de las otras dos.

b) *Escala de actitudes* pretenden evaluar la predisposición u opinión de los sujetos ante determinados temas, tanto referidos a si mismos como a otras personas, que se traducen en una situación física en un determinado punto de la propuesta. Existen tres tipos de escalas, asociadas a sus promotores: Thurstone, Likert y Guttman.

b.1. Escala de **Thurstone**: representa un intento serio para aproximarse a escalas de intervalo, son pues las más perfectas, pero a la vez las más complejas y difíciles de construir, ya que supone que las distancias entre los puntos son similares.

Se parte de un elevado número de enunciados que representan todo el continuo favorable/desfavorable hacia el objeto de la actitud a evaluar y, merced al juicio de los expertos, se asigna a los items seleccionados un determinado valor escalar. Los sujetos eligen entre los items propuestos, siendo su puntuación la suma de los valores de los elegidos. Para la selección de items se siguen los siguientes pasos:

- Recoger gran cantidad de afirmaciones respecto del tema.
- Clasificar estas afirmaciones en once grupos, según las opiniones expresadas por los jueces (50 a 300), desde las más favorables a las más desfavorables.
- Atribuir a cada enunciado un valor escalar igual a la mediana de la asignada por los jueces.
- Elegir un número de enunciados (en torno a 20) repartidos a lo largo de la escala.

En consecuencia la escala resultante se configura sobre una serie de afirmaciones, en torno a las cuáles los sujetos van señalando aquellas con las que están más de acuerdo o los dos o tres items más próximos a su posición personal.

b.2. Escala de **Likert**: consiste en una serie de items sobre los cuáles el sujeto muestra su grado de acuerdo o desacuerdo. Se deben emplear items favorables o desfavorables, pero no los neutrales, pues nos interesa que el sujeto se manifieste de forma expresa. Se trabaja con cinco a siete rangos, aunque en ocasiones se trabaja con un número menor, expresando así la correspondiente gradación en las respuestas.

En realidad se trata de una escala ordinal, dado que permite la clasificación de los individuos en términos de su actitud favorable o desfavorable hacia un determinado objeto o fenómeno. Para alcanzar una correcta interpretación de los resultados globales es preciso determinar con nitidez la validez de dicha escala y los aspectos subyacentes a los items recogidos en el correspondiente instrumento.

b.3. Escala de **Guttman**: su nota diferenciadora consiste en el intento de lograr la unidimensionalidad, esto es, que todos y cada uno de los items estén midiendo la misma

realidad, aunque con distinta intensidad. Por tanto, la elección de un ítem concreto representa el acuerdo con todos los que le preceden en la escala, dada la estrecha relación entre los mismos. La puntuación global consiste en la suma de las puntuaciones en cada elemento con los que está de acuerdo.

En la construcción de cualquier modalidad de escala de entre las propuestas se siguen los siguientes pasos:

- Definición del rasgo y plan previo, consiste en delimitar y acotar la realidad que va a ser evaluada. Es conveniente recurrir en esta fase a consulta a expertos.
- Redacción de los ítems en forma de opiniones con las que se puede estar o no de acuerdo. Por ello es preciso garantizar la relevancia de su contenido, una redacción clara y precisa, incluir ítems discriminativos y bipolares.
- Redacción de las respuestas como alternativas ofrecidas al enunciado de los ítems. Estas se reflejan en forma de letras o números que expresan la gradación del aspecto a evaluar. Entre los diferentes autores unos prefieren el número par (4, 6, 8) para evitar la tendencia a la respuesta central, otros se inclinan por el impar (5, 7, 9).
- Orden y número de ítems, el primer aspecto conviene cuidarlo, pues en ocasiones la respuesta a un ítem condiciona las siguientes cuestiones. En cuanto al número debe ser suficiente para no condicionar la medida del rasgo, si bien, se debe evitar un exceso de ítems que produzca reacciones negativas en el sujeto.
- Análisis y selección de los ítems, para ello sería preciso la realización de un estudio piloto que nos ofrezca datos suficientemente válidos y contrastados sobre el funcionamiento de los mismos a la vez que nos permitirá determinar los niveles de fiabilidad y validez alcanzados.

### **3.2.3. Otros instrumentos para el registro de datos:**

Aunque existen diversos instrumentos para el registro sistemático de los datos en la evaluación de programas educativos, nos vamos a detener de manera más detallada en dos de ellos: el anecdotario y la lista de control.

a) *El anecdotario o incidente crítico*, se trata de un relato descriptivo de episodios o anécdotas de cierta relevancia vinculados a la conducta manifestada por el sujeto observado. En la ficha se deben anotar los incidentes tal y como han ocurrido, sin comentario personal visible, aportando los detalles suficientes que permitan realizar un análisis posterior a jueces imparciales. Su gran valor radica en la riqueza y matización de la información, debidamente circunstanciada y contextualizada. Ha de lograrse consistencia de detalles de una anécdota a la siguiente, con el fin de obtener continuidad en el registro y una posible cuantificación de los resultados. La ficha anecdótica se redacta por buenos

observadores, que sean capaces de ver conductas relevantes para la investigación evaluativa.

b) *La lista de control*, consiste en la enumeración de una serie de conductas cuya presencia o ausencia en el grupo observado se desea constatar. Su objetivo básico consiste en recordarnos que se debe detectar la presencia o ausencia de comportamientos que nos parecen importantes, sin añadir a esta constatación ninguna apreciación cualitativa. En esencia se trata de cuadros de doble entrada en que en un eje se reflejan nombres de los sujetos observados y en el otro la serie de enunciados de conducta analizados, el observador va anotando en el mismo si se da o no esa conducta durante el período de observación. En la construcción se deben recoger las unidades y subunidades de observación debidamente estructuradas, además de ser lo más analíticas posible para garantizar la representatividad de una conducta observable. Como ventaja destaca su gran aplicabilidad por personal no demasiado experto, entre sus inconvenientes sobresale el no diferenciar cuantitativamente ni cualitativamente los diferentes matices que se pueden dar en ese momento.

### **3.3. Consideraciones finales:**

La decisión última sobre la elección del instrumento de recogida de datos, es competencia del propio investigador, que de acuerdo con los objetivos de la evaluación y con las disponibilidades de recursos optará por la elaboración de ese instrumento o el uso de una técnica ya elaborada o bien se decidirá por la construcción total o parcial de dicho instrumento. Sin duda constituye un momento importante en el proceso de la evaluación de programas educativos. Pues la obtención de unos datos rigurosos y válidos garantiza la propia calidad de la investigación evaluativa.

En los momentos iniciales sería más apropiado recurrir al análisis de contenido para valorar la calidad técnica del propio programa en sí mismo. Mientras que la observación va mas ligada a la evaluación del proceso y la entrevista nos sirve tanto en esta parte como en la final sobre los resultados alcanzados. En la misma línea el cuestionario nos sirve en las diferentes fases del proceso, las escalas son apropiadas en la parte final de la evaluación de programas, en cambio las listas de control o anecdotarios son más adecuados para la fase de la valoración procesual o de ejecución del programa.

---

Regresar a [Índice de materias](#)

Regresar a [Contenidos del capítulo](#)

Regresar a [comienzo de la sección](#)

---

## IV. EL ANÁLISIS DE DATOS Y LA ELABORACIÓN DE INFORMES EN LA EVALUACIÓN DE PROGRAMAS EDUCATIVOS:

El método de investigación empleado condiciona la técnica, ya sea ésta de observación, interrogación, análisis o evaluación, lo que nos proporcionará unos datos de diferente naturaleza (cualitativa y / o cuantitativa), que a su vez condicionarán la elección del análisis estadístico seleccionado por el investigador o evaluador.

### 4.1. Importancia del análisis e interpretación de datos:

El análisis de datos recogidos en la evaluación, configura una tarea fundamental, dentro del proceso general anteriormente apuntado, pues nos ayuda a dar sentido y a constatar empíricamente una serie de intuiciones, hipótesis de trabajo, valoraciones y a dar una mayor validez científica a nuestras afirmaciones futuras.

En esencia las fases del análisis de datos se reflejan básicamente en la siguiente propuesta:

- a) Identificar temas y desarrollar conceptos y proposiciones.
- b) Codificar los datos y proceder a su refinamiento.
- c) Relativizar los descubrimientos.

Así pues, las operaciones básicas a desarrollar con los datos se concretan en:

La **depuración**, para evitar la contaminación de los resultados finales. De esta forma se eliminan las respuestas múltiples, el no ajustarse éstas al modelo propuesto. Esta tarea es especialmente importante cuando se va a proceder a los análisis estadísticos mediante tratamiento informático.

La **organización**, para facilitar su lectura, esto se puede llevar a cabo mediante una simple *ordenación* (de mayor a menor, por pesos relativos), lo que nos permite conocer la situación de los sujetos en sus grupos, la *categorización*, estableciendo categorías mediante la diferenciación de grupos en función de las puntuaciones alcanzadas en las diferentes variables contempladas (edad, sexo, nivel socio-económico o de rendimiento, personalidad, integración); la *distribución* de puntuaciones en frecuencias, categorías, grupos. Se trata de conocer las características generales del conjunto de datos, ofreciendo una visión conjunta y sintetizada de los mismos.

La **transformación** que consiste en hacer inteligibles una serie de datos, para que sean asimilables por el evaluador, a la vez que permita su posterior tratamiento estadístico.

#### 4.1.1. Análisis cualitativos y cuantitativos:

El análisis de datos se puede abordar, atendiendo a su naturaleza, desde las metodologías cualitativa o cuantitativa, tal y como hemos visto anteriormente. La primera analiza más los comportamientos que los resultados finales, mientras la segunda se centra más en los valores numéricos de las variables.

En los análisis de tipo **cualitativo** se manifiesta una preponderancia de los datos descriptivos, mediante los cuáles tratamos de proporcionar una descripción de la vida social o académica, pero basándose en un mínimo de interpretación y conceptualización. A través del informe se pretende que los lectores puedan extraer sus propias conclusiones y generalizaciones.

Las fases de estos análisis se concretan en los siguientes momentos interdependientes:

- \* El análisis exploratorio: implica el análisis del contexto con un marco conceptual previo.
- \* La descripción: examen de todos los segmentos en cada categoría, supone una reducción de los datos.
- \* La interpretación: momento arriesgado, pues supone integrar, relacionar, establecer conexiones o comparaciones.

Alcanzar una adecuada interpretación e integración de los datos en el marco de la práctica educativa es fundamental en este tipo de análisis. Ello nos conducirá a establecer nuevas relaciones y consideraciones sobre el fenómeno estudiado. Dado su carácter cíclico nos permite nuevos replanteamientos de forma que repercutan directamente sobre la mejora de actuaciones futuras en el marco escolar.

En los análisis de tipo **cuantitativo**, generalmente más rigurosos y sometidos a mayor control, empleando en su recogida instrumentos más elaborados y con sus correspondientes cálculos de fiabilidad y validez, se suelen seguir las siguientes fases:

- \* Recogida de datos de acuerdo con la especificación de la variable o variables.
- \* Tabulación de los datos.
- \* Representación de distribuciones.
- \* Hipótesis estadísticas de acuerdo con la investigación.
- \* Prueba de hipótesis estadística:
  - Elección de la prueba y nivel de error.
  - Aplicación de la prueba.

- Rechazo o no de la hipótesis planteada.

\* Inferencias respecto a la hipótesis de trabajo.

\* Posible confirmación, ampliación o cambio de la teoría o generalización, propuesta inicialmente.

#### **4.1.2. Cómo se interpretan los datos:**

La interpretación de los datos es, sin duda, un momento importante, tal y como estamos poniendo de relieve, en la investigación evaluativa, pues nos ayuda a comprender el sentido de los mismos dentro del contexto en que actúan.

En la evaluación de programas podríamos establecer una distinción entre lo que es una mera interpretación **aproximativa**, de lo que implica la significación estadística. Así la primera nos ofrece un conocimiento superficial del comportamiento de las variables analizadas, lo que se refiere a si los valores de los grupos son altos, medios, bajos, si se encuentran más o menos concentrados o dispersos, como se distribuyen sobre gráficos, cuál es su grado de relación, etc.

Por su parte la **significación estadística**, viene a añadir criterios de validez científica a la interpretación de los datos. Pues nos permite constatar, con un determinado riesgo de error, si los cambios producidos en el grupo o sujeto, son reales o se pueden explicar por efecto de azar. Ello exige contrastar los valores empíricos de los estadísticos, obtenidos por el investigador, con los valores críticos o de distribución, recogidos en las correspondientes tablas. En el tratamiento informatizado, se nos ofrecen directamente los valores de significación de los estadísticos.

#### **4.1.3. La Estadística y el análisis de datos:**

Podemos afirmar que la Estadística es una ciencia instrumental, de gran ayuda en los análisis de datos; en su acepción científica la debemos considerar como una rama de la Matemática, mientras que desde el punto de vista de la investigación y/o evaluación, prima su carácter instrumental, de apoyo e interpretación de los datos. Desde este enfoque cumple un triple cometido:

\* Facilita el manejo de datos numerosos: función simplificadora y descriptiva.

\* Permite hacer inferencias a la población: hacer extensivos los resultados obtenidos en las muestras.

\* Ayuda en la toma de decisiones: para decidir si procede o no introducir cambios que hayan generado eficacia.

Ahora bien, antes de entrar en los análisis concretos, es conveniente que tengamos presente el tipo de variables que estamos empleando, qué naturaleza tienen; dónde se distinguen, siguiendo la terminología de la presente exposición en: variables cualitativas -categóricas o atributos- y variables cuantitativas ya sean discretas o continuas. Ello va a condicionar el tipo de análisis que debemos efectuar.

#### **4.1.4. Enfoques del análisis de datos:**

En función de las aportaciones de la Estadística, ya reseñadas anteriormente, debemos diferenciar dos niveles en los análisis de datos: descriptivos e inferenciales.

Los análisis **descriptivos** permiten conocer de forma elemental y básica las características de un conjunto de datos o de la situación de un sujeto en un grupo. Podríamos concretar estas funciones en:

- Conocimiento de las características típicas de un conjunto muestral.
- Conocimiento de la relación entre varios conjuntos de datos de variables.
- Predicción del comportamiento en una variable sobre la base de otra u otras.
- Descripción de diferencias y fuentes de variación entre varios conjuntos.

Los análisis de tipo **inferencial**, cumplen una función complementaria a los anteriores, pues nos permiten una visión más global y completa del problema que estamos analizando. El objetivo se centra en derivar conclusiones probabilísticas, partiendo del análisis de los estadísticos muestrales (hacer estimaciones), fijando unos márgenes de error para estas apreciaciones.

En los análisis de tipo cuantitativo centrados en la comprobación empírica de hipótesis, se debe fijar el nivel de significación (margen de error), con que se toma la decisión estadística de aceptar o rechazar las hipótesis previamente formuladas.

#### **4.1.5. Análisis y tratamiento de datos:**

Los datos numéricos recogidos y depurados precisan, para poder ser analizados e interpretados, estar expresados de una forma ordenada y simplificada, ya sea de forma numérica o a través de representaciones gráficas.

En una forma esquemática estas fases, de menor a mayor complejidad estadística, se concretan, según nuestra propuesta en los siguientes puntos:

1. Recopilación y primeros análisis:

- Presentación de los datos: ordenados o no.
- Frecuencias: absolutas, acumuladas, relativas, ajustadas.
- Porcentajes o proporciones.

## 2. Estadísticos de grupo:

- Media.
- Mediana.
- Desviación típica.
- Varianza.

## 3. Representaciones gráficas:

- Datos cualitativos:

- \* Sectores.
- \* Gráficos.
- \* Pictogramas, etc.

- Datos cuantitativos:

- \* Una variable: Diagramas, Histogramas, Polígono de frecuencias, Ojiva.
- \* Dos o más variables: Perfiles (radial, ortogonal), Diagramas de dispersión, Diagramas de desarrollo.

## 4. Situación de un sujeto, dentro de un grupo (puntuaciones individuales):

- Directas.
- Típicas.
- Pentas.
- Eneatipos.
- Cuantiles:
  - \* Cuartiles.
  - \* Centiles o percentiles.

5. Estudio conjunto de dos o más variables:

- Correlación de Pearson.
- Correlación ordinal de Spearman.
- Coeficiente de Contingencia.
- Correlación biserial-puntual.
- Correlación tetracórica.
- Correlación parcial.
- Correlación múltiple.
- Análisis de regresión.

6. Otras metodologías cualitativas:

- Diferencial semántico.
- Metodología "Q".

7. Pruebas para la comprobación empírica de hipótesis:

- Paramétricas:

- \* "t" o "z" (dos grupos independientes o relacionados).
- \* ANAVA (más de dos grupos) Si "F" significativo: Contraste de TUKEY, Contraste de SCHEFFE.

- No paramétricas:

- \* Prueba de Chi cuadrado (nominal).
- \* Prueba "U" de Mann-Whitney (ordinal).
- \* Prueba de Kolmogorov-Smirnov (ordinal).
- \* Prueba "T" de Wilcoxon.
- \* Prueba "H" de Kruskal-Wallis (ordinal).
- \* Prueba "Q" de Cochran (nominal).
- \* Prueba " " de Friedman (ordinal).

8. Técnicas estadísticas multivariadas:

- Para el estudio de la dependencia:

- \* Regresión múltiple.
- \* Análisis discriminante.
- \* Análisis multivariado de la varianza.
- \* Análisis de covarianza.

- Para el estudio de la independencia:

- \* Análisis factorial.
- \* Análisis de componentes principales.
- \* Análisis de correspondencias.
- \* Análisis de "Cluster".

## **4.2. La elaboración de informes en la evaluación de programas:**

El informe de la evaluación de programas educativos debe recoger todo el proceso seguido, con el fin de que lo conozcan las personas implicadas. Consiste pues, en registrar de forma clara y concisa lo que se ha hecho, que caminos se han seguido en cada una de las fases y cuáles son los principales resultados alcanzados.

Desde la perspectiva de implicación pedagógica podríamos distinguir dos tipos de informes: el divulgativo y el científico. En el primero nos centraríamos fundamentalmente en la claridad expositiva y en la comprensibilidad del público a que se dirige, en nuestro caso al profesorado más directamente relacionado con su puesta en marcha, por ello le interesan especialmente todos los aspectos relacionados con las implicaciones que se derivan del mismo. El segundo nivel del informe debe prestar especial atención a los aspectos relacionados con el rigor científico y metodológico de los planteamientos desarrollados. Ello puede dar origen a publicaciones, informes técnicos dirigidos a las autoridades u organismos que subvencionan los proyectos.

En líneas generales en el informe de la evaluación de programas se deben reflejar los siguientes puntos:

1. Introducción y marco general dónde se inscribe el proyecto, ello nos permitirá conocer en que ámbito nos movemos y a qué público nos dirigimos. Se trata de reflejar aquellos aspectos contextuales que tanta importancia tienen en este tipo de trabajos.
2. Definición clara y concisa del problema que es objeto de investigación y análisis.
3. Análisis del estado de la cuestión, mediante la consulta a diversas fuentes, ya sean de tipo bibliográfico o no, de todos los aspectos relacionados con el problema que nos ocupa, como se ha abordado en otros estudios, con qué resultados.

4. Formulación de las hipótesis de investigación o de los objetivos de la misma, según que el enfoque se encuadre en la metodología cuantitativa o cualitativa. Ello va a marcar todo el proceso práctico de ejecución.

5. Enfoque metodológico propuesto, conforma toda la parte operativa del proceso de evaluación de programas. En esta parte se deben recoger los aspectos relacionados con dicho enfoque, haciendo especial hincapié entre otros posibles a:

- Las variables que intervienen y su definición operativa en el supuesto concreto que analizamos.

- Los sujetos que van a intervenir en la evaluación, así como los cometidos a cubrir. También en este momento se debe reflejar, cuando así se produzca, los procesos de selección de la muestra así como de su composición.

- En las propuestas identificadas con el paradigma cuantitativo es conveniente matizar los diferentes aspectos relacionados con el control de variables extrañas.

- Los instrumentos de recogida de datos, dónde caben dos opciones: bien que se trate de la selección de alguno ya existente, para lo cuál es preciso tener en cuenta una serie de recomendaciones específicas; bien sea necesario proceder a la elaboración de uno nuevo, lo que lleva consigo una tarea laboriosa y con rigor. En muchas ocasiones las aportaciones básicas de la investigación evaluativa radican precisamente en la confección de los correspondientes instrumentos de recogida de datos.

- La recogida de datos, consiste en acudir a las fuentes primarias productoras de los mismos.

- El análisis de los datos va ligado a su presentación, siempre tratando de responder a los objetivos de la investigación. En realidad se van a reflejar las novedades aportadas en cada uno de los casos. En el supuesto de la investigación cuantitativa el análisis de datos concluye con la decisión estadística de la aceptación o el rechazo de las hipótesis previamente formuladas. Es importante que la presentación de los análisis sea comprensible para el público a que se dirige el estudio.

6. Conclusiones y recomendaciones que se pueden derivar del trabajo evaluativo, sin duda configuran la parte más auténticamente pedagógica y ligada a la actividad práctica. Del análisis de los datos se derivarán unas conclusiones que a su vez pueden y deben generar un conjunto de recomendaciones especialmente útiles para amplios colectivos. Si bien, conviene ser cautos en la extrapolación de resultados a colectivos más amplios.

La redacción del informe debe recoger una crítica, desde argumentos científicos al propio trabajo evaluativo realizado, así como las posibilidades y limitaciones que de ello se pueden derivar y los campos que nos puede abrir para futuras profundizaciones específicas. Todo ello debe contribuir a que las aportaciones no se queden en una mera recopilación de textos

y datos, sino que trasciendan la propia realidad del fenómeno estudiado y puedan ser de utilidad para otros colectivos interesados en el tema.

---

Regresar a [Índice de materias](#)

Regresar a [Contenidos del capítulo](#)

Regresar a [comienzo de la sección](#)

---

## BIBLIOGRAFÍA

ABARCA PONCE, M.P.: **La evaluación de programas educativos**. Ed. Escuela Española. Madrid, 1989.

ADAMS, G.S.: **Medición y evaluación en Educación, Psicología y "Guidance"**. Ed. Herder. Barcelona, 1983.

ANGUERA, M.T.: **Metodología de la observación en las Ciencias Humanas**. Ed. Cátedra. Barcelona, 1985.

BLOOM, B. y otros.: **Evaluación del aprendizaje**. Ed. Troquel. Buenos Aires, 1981.

CABRERA, F.A. y ESPIN, J.V.: **Medición y evaluación educativa**. Ed. PPU. Barcelona, 1986.

CARDONA, J. (Coord.): **Metodología innovadora de evaluación de centros educativos**. Ed. Sanz y Torres. Madrid, 1994.

COHEN, L. y MANION, L.: **Métodos de investigación educativa**. Ed. La Muralla. Madrid, 1990.

COOK, T.D. y REICHARDT, Ch.S.: **Métodos cualitativos y cuantitativos en investigación evaluativa**. Ed. Morata. Madrid, 1986.

CORONA, R. y otros.: **Manual de técnicas de evaluación y ajuste de información estadística**. Ed. Fondo de Cultura Económica. México, 1982.

CRONBACH, L.J.: **Designing evaluations of educational and social programs**. Ed. Jossey-Bass. San Francisco. USA, 1982.

DE KETELE, J.M.: **Observar para educar**. Ed. Visor. Madrid, 1984.

DE LANDSHEERE, G.: **Diccionario de la evaluación y de la investigación**. Ed. Oikos-tau. Barcelona, 1985.

DELORME, Ch.: **De la animación pedagógica a la investigación-acción**. Ed. Narcea. Madrid, 1985.

DEMO, P.: **Investigación participante**. Ed. Kapelusz. Buenos Aires, 1985.

DENDALUCE, I. (Coord.): **Aspectos metodológicos de la investigación educativa**. Ed. Narcea. Madrid, 1988.

EBEL, R.L.: **Fundamentos de la medición educacional**. Ed. Guadalupe. Buenos Aires, 1977.

----- **La investigación-acción en el aula**. Ed. Generalitat Valenciana. Valencia, 1986.

ELLIOT, J.: **La investigación-acción en Educación**. Ed. Morata. Madrid, 1989.

ETXEBERRIA, J. y otros: **Análisis de datos y textos**. Ed. Ra-Ma. Madrid, 1995.

FERNÁNDEZ PÉREZ, M.: **Evaluación y cambio educativo. Análisis cualitativo del fracaso escolar**. Ed. Morata. Madrid, 1988.

GARCÍA HOZ, V. y PÉREZ JUSTE, R.: **La investigación del profesor en el aula**. Ed. Escuela Española. Madrid, 1984.

GARCÍA HOZ, V. (Coord.): **Problemas y métodos de investigación en educación personalizada**. Ed. Rialp. Madrid, 1994.

GARCÍA LLAMAS, J.L.: **El aprendizaje adulto en un sistema abierto y a distancia**. Ed. Narcea. Madrid, 1986.

GIMENO, J. y PÉREZ, A.: **La enseñanza: su teoría y su práctica**. Ed. Akal. Madrid, 1983.

GOETZ, J.P. y LE COMPTE, M.D.: **Etnografía y diseño cualitativo en la investigación educativa**. Ed. Morata. Madrid, 1988.

GOYETTE, G. y LESSARD-HEBERT, M.: **La investigación-acción. Funciones, fundamentos e instrumentación**. Ed. Laertes. Barcelona, 1988.

GUBA, E.G.: **The paradigm dialog**. Ed. Sage. Newbury, 1990.

GUBA, E.G. and LINCOLN, Y.S.: **Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches**. Ed. Jossey-Bass. San Francisco, 1992.

HILLS, J.R.: **Evaluación y medición en la escuela**. Ed. Kapelusz. México, 1981.

HUARTE, F. (Coord.): **Temas actuales sobre Psicopedagogía y Didáctica**. Ed. Narcea. Madrid, 1988.

KEMMIS, S. y McTAGGART, R.: **Cómo planificar la investigación-acción**. Ed. Laertes. Barcelona, 1988.

KERLINGER, F.N.: **Investigación del comportamiento. Técnicas y metodología**. Ed. Interamericana. México, 1981.

LAFOURCADE, P.D.: **La evaluación en organizaciones educativas centradas en logros**. Ed. Trillas. México, 1982.

LEWY, A.: **Manual de evaluación formativa del currículo**. Ed. UNESCO. Madrid, 1976.

MEHRENS, W. y LEHMANN, I.: **Medición y evaluación en la Educación y en la Psicología**. Ed. CECSA. México, 1982.

MIGUEL, M. de; MORA, J.G. y RODRÍGUEZ, S. (Coords.): **La evaluación de instituciones universitarias**. Ed. Consejo de Universidades. Madrid, 1991.

MILES, M.B. and HUBERMAN, A.M.: **Qualitative data analysis: A sourcebook of new methods**. Ed. Sage. Beverly Hills, 1984.

MORALES, P.: **Medición de actitudes en Psicología y Educación**. Ed. Trartralo. San Sebastián, 1988.

ORDEN HOZ, A. de la: **La evaluación educativa**. Ed. Docencia. Buenos Aires, 1982.

PASCUAL, R. (Coord.): **La gestión educativa ante la innovación y el cambio**. Ed. Narcea. Madrid, 1988.

PÉREZ JUSTE, R. y GARCÍA RAMOS, J.: **Diagnóstico, evaluación y toma de decisiones**. Ed. Rialp. Madrid, 1989.

PÉREZ JUSTE, R. y MARTÍNEZ ARAGÓN, L.: **Evaluación de centros y calidad educativa**. Ed. Cincel. Madrid, 1989.

PÉREZ JUSTE, R.: **El modelo evaluativo en la Reforma**. Ed. UNED. Madrid, 1993.

PÉREZ SERRANO, G.: **Investigación cualitativa. Retos e interrogantes.** I y II. Ed. La Muralla. Madrid, 1994.

POPHAM, W.J.: **Problemas y técnicas de la evaluación educativa.** Ed. Anaya. Madrid, 1980.

----- **Evaluación basada en criterios.** Ed. Magisterio Español. Madrid, 1983.

RODRÍGUEZ SAN MARTÍN, A.: **Fundamentos y estructura de la evaluación educativa.** Ed. Anaya. Madrid, 1983.

ROSALES, C.: **Criterios de una evaluación formativa.** Ed. Narcea. Madrid, 1988.

ROSSI, P. y FREEMAN, H.: **Evaluación. Un enfoque sistémico para programas sociales.** Ed. Trillas. México, 1989.

SELLTIZ, C. y otros: **Métodos de investigación en las relaciones sociales.** Ed. Rialp. Madrid, 1980.

SCRIVEN, M.: **The logic of evaluation.** Ed. Edgepress. Inverness, 1980.

----- **Evaluation Thesaurus.** Ed. Sage. Newbury Park, 1991.

SHADISH, W.R.Jr.; COOK, T.D. and LEVITON, L.C.: **Foundations of program evaluation. Theories of practice.** Ed. Sage. Newbury Park, 1991.

STAKE, R.E.: **Program evaluation, particularly responsive evaluation.** Ocasional Paper Series nº 5. University of Western Michigan, 1975.

----- **Evaluating the arts in education: An responsive approach.** Ed. Merrill. Columbus, 1975.

STHENHOUSE, L.: **La investigación como base de la enseñanza.** Ed. Morata. Madrid, 1987.

STUFFLEBEAM, D.L. y SHINKFIELD, A.J.: **Evaluación sistemática: guía teórica y práctica.** Ed. Paidós-MEC. Madrid, 1987.

TAYLOR, S.J. y BOGDAN, R.: **Introducción a los métodos cualitativos de investigación.** Ed. Paidós. Buenos Aires, 1986.

TENBRINK, T.D.: **Evaluación. Guía práctica para profesores.** Ed. Narcea. Madrid, 1984.

THORNDIKE, R.L. y HAGEN, E.: **Tests y técnicas de medición en Psicología y Educación.** Ed. Trillas. México, 1978.

----- **Medición y evaluación en Psicología y Educación.** Ed. Trillas. México, 1989.

TYLER, R.E.: **Principios básicos del currículo.** Ed. Troquel. Buenos Aires, 1982.

VALETT, R.E.: **Evaluación del desarrollo de aptitudes básicas para el aprendizaje.** Ed. TEA. Madrid, 1983.

VILLAR ANGULO, L.M.: **Evaluación diagnóstica de los procesos mentales de los profesores.** Ed. Universidad de Sevilla, 1987.

WALKER, R.: **Métodos de investigación para el profesorado.** Ed. Morata. Madrid, 1989.

WARREN, C.: **Qualitative Research methods.** Ed. Sage. London, 1988.

WEIS, C.H.: **Investigación evaluativa.** Ed. Trillas. México, 1975.

WOODS, P.: **La escuela por dentro: la etnografía en la instrucción educativa.** Ed. Paidós-MEC. Madrid, 1987.

WITTROCK, M.C.: **La investigación de la enseñanza I y II.** Ed. Paidós-MEC. Barcelona, 1989, 1990.

ZABALZA, M.A.: **Diseño y desarrollo curricular.** Ed. Narcea. Madrid, 1987.