

UNIVERSIDAD PEDAGOGICA NACIONAL

LICENCIATURA EN EDUCACION

PLAN 1994

APLICACION DE LA ALTERNATIVA DE INNOVACION

- ANTOLOGIA -

JUNIO DE 1995.

000

CONTENIDO

I	PRESENTACION	
II	PRIMERA PARTE	
	"El manejo de la ansiedad en la investigación", George Devereux	50
	"La evaluación", Wheeler	40
	"La evaluación", Fernando Cembranos	21
	"Recopilación de datos", Merlin C: Wittrock.	320
III	SEGUNDA PARTE	51
	"Estrategias de evaluación en el aula", C.E.T.E.	59
	"Tests hechos por el profesor", Louis J. Karmel.	60
	"Ficha para ingreso de los alumnos", Ma. A. Galindo.	90
	"Recolección de datos", Roberto Hernández y cols.	100
IV	BIBLIOGRAFIA GENERAL	107
		209

I PRESENTACION

El curso "Aplicación de la alternativa de innovación" tiene un carácter eminentemente operativo dentro del plan de estudios, su propósito es brindar un espacio para que el profesor-alumno realice la puesta en práctica de la alternativa que haya elaborado en el curso anterior; en esta sentido, al desarrollo del curso se centra en los materiales y productos que los propios profesores-alumnos elaboran, la recuperación de textos constituye una actividad complementaria y determinada por las necesidades particulares de cada estudiante; en este sentido se puede decir que no se requiere de una bibliografía básica, sino que todos los textos son complementarios. Por ello para este curso se decidió presentar solo una antología; los textos de ésta se han dividido en dos apartados, el primero que incluye textos generales y cuya revisión se recomienda para todos los participantes del curso, y el segundo se puede considerar como un anexo.

Los artículos incluidos en la primera parte ofrecen planteamientos conceptuales amplios que pueden ser recuperados para orientar el proceso que se seguirá en este curso.

Instrumentos de la Licenciatura
En la segunda parte se ubican seis textos que se han seleccionado por ofrecer elementos específicos para la elaboración y manejo de distintos instrumentos que se usan para la recuperación de información, su revisión depende del tipo de instrumentos o estrategias que el profesor-alumno haya seleccionado, además de que se constituyen en un banco de información al que se puede recurrir en caso necesario; se le recuerda que en distintos cursos de la licenciatura se ofrece información sobre algunos instrumentos para la recuperación de información, en particular para aquellos trabajos relacionados con la gestión escolar se sugiere revisar el curso de "evaluación y seguimiento en la escuela".

Por la naturaleza propia de los textos presentados y su papel dentro de las actividades del curso no se consideró necesario hacer presentaciones demasiado puntuales. Cada uno de ellos, por ello la presentación de los textos se hace de manera muy general dentro de la presentación del apartado correspondiente y no al principio de cada lectura como sucede en las antologías de los demás cursos de la licenciatura.

II PRIMERA PARTE

Los textos que se incluyen en esta primera parte de la antología se han seleccionado con intenciones distintas, el primero de ellos tiene la finalidad de hacerle reflexionar sobre el desempeño que tienen las emociones de quien investiga en el proceso de investigación y en particular en la interpretación de la información; los siguientes tres textos tienen un carácter distinto, ofrecen conceptos e información sobre las actividades principales que desarrollará en este curso. Las lecturas no se presentan en una secuencia que defina previamente su revisión, ni se espera que dentro del curso usted realice un análisis puntual de los mismos; su recuperación depende de usted y de las orientaciones que le de su asesor.

El texto de Devereux que se incluye en esta antología es de corte psicoanalítico, recupera los conceptos de transferencia y contratransferencia manejados por Freud para identificar la forma en que un investigador se ve involucrado emocionalmente en el proceso de investigación y como la ansiedad generada puede volverse un obstáculo que impida o distorciona el proceso general de la investigación; señala los mecanismos de resistencia que comúnmente se observan en los investigadores y describe algunas estrategias que pueden servir para superar las dificultades mencionadas. En esta antología se incluye el apartado referido a la argumentación inicial así como los capítulos I y VII que son los más significativos para el proceso que se aborda en este curso. Le sugerimos que realice su lectura antes de presentarse a la primera sesión de trabajo con su asesor o grupo.

En segundo lugar se le presenta el texto "La evaluación" de Wheeler el cual aborda este tema ubicándolo dentro del proceso de desarrollo del currículum escolar, el autor ofrece elementos que permiten delimitar y conceptualizar a la evaluación y como se debe concretizar en aspectos específicos.

El texto de Cembranos, que aparece en tercer término, complementa la información proporcionada por el anterior, realiza el análisis de algunos términos relacionados con la evaluación que no son abordados en el texto de Wheeler como son: la flexibilidad, la sensibilidad social, la creatividad, etc. Además ofrece orientaciones útiles para la selección de instrumentos para la recuperación de información en función del tipo y objetivo de la evaluación que se pretenda realizar. La revisión de estos dos textos le brindarán elementos importantes para sus actividades.

Por último se presenta un texto de Wittrock que hace referencia explícita al proceso de recuperación de información dentro de la investigación, seguramente su revisión resultará altamente significativa ya que el autor expone una serie de consideraciones que son determinantes para el proceso mencionado, que es una actividad central del curso, y consecuentemente para la evaluación posterior. Revise este texto antes de iniciar la segunda fase del curso para que su lectura sea mas provechosa.

LA ARGUMENTACIÓN

El punto de partida de mi obra es una de las proposiciones más fundamentales de Freud, modificada a la luz de la concepción einsteiniana de la fuente de los datos científicos. Decía Freud que la transferencia es el dato más fundamental del psicoanálisis considerado como método de investigación. A la luz de la opinión de Einstein de que sólo podemos observar los acontecimientos "en" el observador —de que sólo sabemos lo que sucede en y al aparato experimental, cuyo componente más importante es el observador— he ido un paso más allá por el camino que dejara Freud. Afirmo que es la contratransferencia y no la transferencia el dato de importancia más decisiva en toda la ciencia del comportamiento. porque la información que se puede sacar de la transferencia por lo general también puede obtenerse por otros medios, y no sucede así con la que proporciona el análisis de la contratransferencia. Es válida esta especificación, aunque transferencia y contratransferencia sean fenómenos conjugados igualmente básicos; sencillamente porque el análisis de la contratransferencia es científicamente más productivo en datos acerca de la naturaleza del hombre.

El estudio científico del hombre

- 1] es impedido por la ansiedad que suscita traslape entre sujeto y observador,
- 2] que requiere un análisis de la naturaleza y el lugar donde se deslindan ambos;
- 3] debe compensar lo parcial de la comunicación entre sujeto y observador en el nivel consciente; pero
- 4] debe rehuir la tentación de compensar la integridad de la comunicación entre sujeto y observador en el nivel inconsciente;
- 5] que causa ansiedad y por ende reacciones contratransferenciales;
- 6] deforma la percepción e interpretación de los datos, y
- 7] produce resistencias contratransferenciales que se disfrazan de metodología, lo que ocasiona nuevas distorsiones *sui generis*.
- 8] Puesto que la existencia del observador, sus actividades obser-

¹ En esta perspicaz reseña de mi relato acerca de la psicoterapia de un niño de los llanos (1951a). Caudill (1951) destaca que yo había registrado pero no analizado mis reacciones de contratransferencia. La omisión era intencional, porque todavía no escribía yo esta obra.

DEVEREUX, George: "La argumentación" en: De la ansiedad al método en las ciencias del comportamiento. México, Siglo XXI, 1992, pp. 3-32.

vacionales y sus angustias (aun en la observación de sí mismo) producen distorsiones que son no sólo técnica sino también lógicamente imposibles de eliminar.

9] toda metodología efectiva de la ciencia del comportamiento ha de tratar esos trastornos como los datos más significativos y característicos de la investigación de la ciencia del comportamiento, y

10] debe usar la subjetividad propia de toda observación como camino real hacia una objetividad auténtica, no ficticia,

11] que debe definirse en función de lo realmente posible y no de lo que "debería ser".

12] Si se pasan por alto o se desvían por medio de resistencias contratransferenciales disfrazadas de metodología, esos "trastornos" se convierten en fuentes de error incontroladas e incontrolables, mientras que

13] si se tratan como datos básicos y característicos de la ciencia del comportamiento, son más válidos y productores de *insights* que cualquier otro tipo de datos.

En resúmenes cuentas, los datos de la ciencia del comportamiento suscitan ansiedades: las que se trata de eludir por una pseudometodología inspirada por la contratransferencia; esta manobra es la causante de casi todos los defectos de la ciencia del comportamiento.

El gran matemático Lagrange dijo hace mucho tiempo que la naturaleza sencillamente no hace caso de las dificultades que plantea a los científicos, cuya misión — como declaró en otra ocasión — consiste en buscar la simplicidad, pero también en desconfiar de ella. Quiere esto decir que el mejor — y quizá el único — medio de alcanzar una simplicidad congruente con los hechos, es lidiar de frente con las mayores complejidades, mediante el artificio, extremadamente práctico de tratar la dificultad per se como un dato fundamental, que no debe rebuizarse sino aprovecharse al máximo en experiencia más simples.

Los primeros capítulos de este libro, que destacan la angustia causada por los datos de la ciencia del comportamiento, pueden dar la impresión errónea de que la objetividad es imposible a priori en la investigación de la ciencia del comportamiento y que para reducir al mínimo las deformaciones debidas a la subjetividad debemos interponer más y más filtros — tests, técnicas de entrevista, accesorios y otros artificios heurísticos — entre nosotros y nuestros sujetos. Incluso podría parecer que el mejor "observador" es una máquina, y que el observador humano debería aspirar a una suerte de invisibilidad que — si se lograra — eliminaría al observador de la situación observacional.

Este modo de enfocar olvida implícitamente que cada uno de esos filtros, al mismo tiempo que "corrige" algunas distorsiones debidas a la subjetividad, produce otras deformaciones específicas propias, por lo general inadvertidas. Y por encima de todo, cierra que en un sentido aristotélico ("Parva Naturalia", 455a, 10^o, 11, incluso el observador invisible tendrá que decir: "Y esto percibo, yo" ... independientemente de que aquello que percibe sea el comportamiento mismo, un electroencefalograma o un resultado matemático (capítulo XXII). Además, en algún punto ha de decir también: "Esto significa que..." Esto es técnicamente una "decisión" y es un hecho fundamental el que la "teoría de los juegos" no puede producir decisiones, sino sólo definir consecuencias y calcular su probabilidad. La decisión — que en la ciencia consiste en decir "Esto significa que..." — la sigue tomando el científico del comportamiento, de acuerdo con la misma subjetividad y en respuesta a las mismas angustias a que se enfrenta cuando no emplea ninguna clase de filtros. Por eso yo no preconizo la eliminación de los filtros y sólo insisto en la eliminación de la ilusión de que suprimen toda subjetividad y neutralizan por completo la angustia. Y no es así; tan sólo splazan ligeramente el lugar de destino entre sujeto (objeto) y observador y postponen el momento exacto en que aparece el elemento subjetivo (decisión). Una cosa es elegir el lugar del destino y el "momento de la decisión", en que el hecho se transforma en verdad de modo óptimo, y otra cosa es pretender que al hacerlo, así suprimimos toda angustia y subjetividad. Además, aun cuando se elijan óptimamente ese lugar y ese momento, todavía habrá que tomar en cuenta las deformaciones producidas por los filtros, las manipulaciones y otros artificios que hacen posible esta disposición "óptima".

No se hace buena ciencia pasando por alto sus datos, más fundamentales y característicos que son, muy concretamente, las dificultades propias de esa ciencia. El científico del comportamiento no puede ignorar la acción recíproca de sujeto y objeto: sea la esperanza de que, si durante bastante tiempo hace como que no existe, acabará por desaparecer bonitamente.

El negarse a utilizar estas dificultades creativamente sólo puede llevar a la recolección de datos cada vez menos pertinentes, más segmentarios, periféricos y aun triviales, que no derramarán ninguna luz sobre lo que hay de vivo en el organismo o de humano en el hombre. Por eso el científico debe cesar de destacar exclusivamente su manipulación del sujeto y tratar de entender al mismo tiempo — y a veces primordialmente — a sí mismo *qua* observador. En este sentido, todo experimento con ratas es también un experimento.

rimiento hecho en el observador, cuyas angustias y maniobras para rehuir el cuerpo, tanto como su estrategia de investigación, percepción de datos y toma de decisiones (interpretación de los datos) pueden derramar más luz sobre la naturaleza del comportamiento en general de lo que podría hacerlo la observación de las ratas —y aun de otros seres humanos.

Esto implica que las dificultades tradicionales de la ciencia del comportamiento no se deben sólo a una determinación poco juiciosa del lugar y la naturaleza del deslinde entre datos "reales" y productos "incidentales" o epifenomenales de la estrategia de investigación. Indica que el sujeto más capaz de manifestar un comportamiento científicamente-utilizable es el mismo observador. Esto significa que un experimento con ratas, una excursión antropológica o un psicoanálisis contribuyen más a la comprensión del comportamiento si se ven como fuente de información acerca del psicólogo de los animales, el antropólogo o el psicoanalista que si se consideran tan sólo una fuente de información acerca de las ratas, los primitivos o los pacientes. En una verdadera ciencia del comportamiento, los primeros datos son básicos; los otros son epifenomenales... o sea, habiendo en puridad, subproductos que, naturalmente, también ameritarían su aprovechamiento.

No es el estudio del sujeto sino el del observador el que nos proporciona acceso a la esencia de la situación observacional.

1. El comportamiento del "comportamiento" son síntomas triples:

2. Los "rasgos" producidos por la existencia y las actividades observacionales del observador.

3. El comportamiento del observador: sus angustias, sus maniobras defensivas, su estrategia de investigación, sus "decisiones" (= su atribución de un significado a lo observado).

Por desgracia, es de este tercer tipo de comportamiento del que tenemos menos información, porque nos hemos negado sistemáticamente a estudiar la realidad como ella lo requiere. Muchos de los datos que claré son entonces consecuencia de mis intentos de entender mi propio comportamiento, tanto de etólogo de campo como de psicoanalista clínico, completados por los *insights* que haya podido sacar del estudio de Lévi-Strauss (*Tristes Tropiques*, 1955), de Bandier (*Afrique ambiguë*, 1957), y de Condominas (*L'Exotisme est quotidien*, 1963), que son los únicos intentos de talla, que yo conozca, de apreciar el impacto causado en el científico por sus datos y su actividad científica. Y ciertamente, por grande que sea la importancia de la obra *objetive* de Lévi-Strauss, la importancia de sus *Tristes Tropiques* aún puede ser

mayor para el futuro de la ciencia del comportamiento... debido en parte a que también aumenta nuestra penetración de los datos objetivos y los descubrimientos de Lévi-Strauss.

Como hubiera sido impropio analizar estas tres obras autobiográficas extremadamente sintéticas en profundidad, me veo obligado a citar principalmente mis propias observaciones, completadas por muchos ejemplos breves del comportamiento de otros científicos que no han considerado necesario escudriñarse a sí mismos. La ciencia no pierde nada con ello, ya que el análisis de un gran número de hechos relativamente superficiales —que ilustren la *gama* o amplitud de los fenómenos— proporciona exactamente los *mis insights* que el análisis en profundidad de un solo fenómeno (Devereux, 1955a). Amplitud es profundidad, rotando 90° en posición horizontal: la profundidad es amplitud si los 90° giran para ocupar una posición vertical. La equivalencia de ambos radica en la hipótesis ergódica. Tomado por separado, cada uno de mis "casos" es una anécdota; todos juntos, son un análisis de la amplitud —y por lo tanto en profundidad— de las reacciones del científico a sus datos y su "hacer ciencia".

Los últimos capítulos de este libro muestran cómo se emplean a manera de puentes las situaciones que suelen considerarse barreras.

4. La ciencia del comportamiento se volverá simple cuando empiece a tratar las propias reacciones del científico a su material y su trabajo como el más fundamental de todos los datos de esa ciencia. Mientras tanto, sólo tendremos la ilusión de la simplicidad.

Suele decirse de las obras dedicadas a los seres humanos que son duras o blandas de ánimo. La mía no es de las unas ni de las otras, puesto que aspira a la objetividad en relación con esa ternura que hace imposible toda ciencia realista del comportamiento. Cualquier libro que trata del hombre tiene un interés humano, y esto debe dejarse bien sentido. No creo que el Hombre necesite salvarse de sí mismo; le basta con ser él mismo. El siglo veintiuno más de los hombres que de los "humanistas". La Grecia del siglo V fue sencillamente humana; se hizo "humanista" en reacción a los horrores de la guerra del Peloponeso. Esquilo, el combatiente

La hipótesis ergódica postula que los mismos resultados pueden obtenerse haciendo un número infinito de monedas simultáneamente o una sola moneda un número infinito de veces.
"Tough minded or tender minded" en el original. Cargorfes utilizadas por el filósofo norteamericano William James para caracterizar dos tipos de experiencia científica. [art.]

de Maratón y poeta de las *Euménides*, no fue un humanista. Sócrates, figura de transición, fue más que un humanista. Platón fue un humanista, puesto que en nombre de la humanidad trató de salvar a ésta de sí misma. Toda filosofía de opresión procede de la de Platón (Popper, 1962), cuya filantropía era desdén, porque trataba al hombre como un objeto de contemplación y manipulación. En este sentido, el científico del comportamiento que se autocalifica de "duro" es un filántropo desdénoso: un "humanista" espurio. La ciencia del comportamiento auténtica nacerá cuando quienes la practiquen comprendan que una ciencia realista del género humano sólo pueden crearla hombres perfectamente conscientes de su propia humanidad, precisamente cuando más plenamente la pongan por obra en su labor científica.

CAPÍTULO I

EN BUSCA DE UNA CIENCIA CIENTÍFICA DEL COMPORTAMIENTO

El orden en que las nociones del hombre acerca de los diversos segmentos de realidad se hicieron científicas fue determinado en gran parte por su mayor o menor interés afectivo en los diversos campos de fenómenos. Cuanto mayor ansiedad ocasiona un fenómeno, menos capaz parece el hombre de observarlos debidamente, de pensarlo objetivamente y de crear métodos adecuados para describirlo, entenderlo, controlarlo y pronosticarlo. No es casualidad que los tres hombres que más radicalmente modificaron nuestro concepto del hombre en el universo —Copérnico, Darwin y Freud— nacieran en este orden. Era más fácil ser objetivo acerca de los cuerpos celestes que del hombre en tanto que organismo, y esto más fácil que la objetividad acerca de la personalidad y el comportamiento del hombre. Si Freud hubiera sido contemporáneo de Copérnico, y aun de Darwin, no hubiera podido crear un concepto psicoanalítico del hombre aunque hubiera habido los medios para recoger y comparar los datos brutos necesarios y él hubiera tenido acceso a ellos; y muchos de ellos habían ya sido empleados correctamente, aunque para fines no científicos, por los brujos primitivos (Devereux, 1961a). Por cierto que lo más nuevo del psicoanálisis no es su teoría sino la posición metodológica según la cual la tarea principal de la ciencia del comportamiento es el análisis de la idea que el hombre tiene de sí mismo. Esta actitud revolucionaria sólo se hizo psicológicamente tolerable después de haber Copérnico y Darwin revaluado el lugar del hombre en el cosmos y en el plan de la vida.

Es un hecho histórico —si bien, como espero demostrar, no una necesidad inevitable— que el interés afectivo del hombre por los fenómenos que estudia con frecuencia le impide ser objetivo en relación con ellos.

La primera de las tres grandes revoluciones científicas fue la de Copérnico, sencillamente porque el hombre no es muy sentimental en materia de astronomía. Por paradójico que pueda parecer, la mejor prueba de esto la constituyen los mitos astrales. Al transferir

¹ Empleo la palabra filantropía en su sentido moderno y no en el de la Poética de Aristóteles, 1452b33.

Los conflictos interiores e interpersonales causantes de angustia a la bóveda de los cielos, el hombre logró disociarse de los problemas que lo asediaban y especular acerca de ellos con cierta objetividad. Así como Zeus solía desahacerse de los personajes míticos molestos poniéndolos entre las estrellas, así hoy, cuando el psicólogo de las ratas no sabe qué hacer y está en el umbral de sus emociones, su imaginación hace de la rata de carne y hueso un modelo casi platónico de "sura rat".

El hecho de que el interés emocional del hombre por un fenómeno dado pueda ser inversamente proporcional a su objetividad para con él puede observarse mejor entre los primitivos... no por que sean incapaces de objetividad, sino porque no suelen darse cuenta de su falta de objetividad.

Caso 1: Aunque los mois sedang tienen muchas ideas super-naturalistas y erróneas acerca de todos los animales, en conjunto el grado en que su "historia natural" es "historia innatural" depende de la importancia del papel que esa especie animal desempeña en su vida. Y así, sus ideas acerca de los tigres son más complejas en super-naturalistas que sus nociones acerca de las ratas de la selva, y tienen creencias más poco realistas de los carabaos, puercos y perros que de los simples pollos o gallinas. De modo semejante, todos los hombres tienen ideas más tontas acerca de la sexualidad que, digamos, de la comida, sencillamente porque el hombre se emociona más con lo sexual que con el alimento.

La ciencia del comportamiento es asimismo menos científica que la física o la biología porque a los fenómenos físicos los determina un pequeño número de variables relativamente fáciles de cuantificar, mientras que el comportamiento del hombre puede engendrarse sólo en función de un número muy grande de variables. Además, un conocimiento razonablemente completo del estado de un sistema físico en el tiempo $t + \Delta t$, mientras que para predecir su estado en el tiempo $t + \Delta t$, mientras que para predecir el comportamiento del hombre en el tiempo $t + \Delta t$ debemos conocer su estado no sólo en el momento precedente, t , sino también durante toda su vida, puesto que el hombre es un sistema cronohistórico, cuyo comportamiento es determinado más profundamente por un tipo de "memoria" algo parecido a la histeresis en física,¹ que por su estado y situación actuales.

¹ Rata estadística (K.T.)

² Sistemáticamente, esto significa que el comportamiento de muchos sistemas físicos puede describirse por medio de ecuaciones diferenciales, mientras que el del hombre sólo puede describirse por medio de ecuaciones integrodiferenciales.

En resumen, las ciencias del comportamiento son ahora ciencias científicas que las físicas debido a:

1) El mayor interés emocional del hombre por el género humano, y por sí mismo que por los objetos materiales;

2) La inherente complejidad del comportamiento y la necesidad de entenderlo cronohistóricamente.

Los científicos del comportamiento, embarazados por el hecho de que su disciplina va a la zaga de la ciencia física, tratan de compensarlo limitando los procedimientos de la física. Hay algunos incluso que sólo estudian los fenómenos cuantificables, y por el momento no hacen caso de todos los datos—por muy conspicuos e importantes que sean—que no son fácilmente cuantificables. Al hacerlo así, implícitamente dejan de distinguir entre las técnicas de la física, determinadas primordialmente por la distinta naturaleza de los fenómenos físicos, y el método científico general, que tiene una validez interdisciplinaria y por ello es igualmente aplicable a la ciencia física y al comportamiento. Esta distinción es lógicamente legítima aunque, hablando históricamente, la mayoría de las reglas del método científico se formularon de acuerdo con los procedimientos de la ciencia física.² Por desgracia, la transposición mecánica de las técnicas de la ciencia física a las demás ciencias—como por ejemplo la cuantificación observativa—puede llevar a la falacia lógica de que la mera cuantificación hace automática mente científico un dato.

Caso 2: El tercer Kinsey Report (Gebhard et al., 1958) "demuestra" estadísticamente que el aborto no es traumático. Sea esto falso o verdadero, no es un enunciado científico, a pesar—y casi a causa—de que el hecho está "sustentado" por estadísticas. Por cierto que si bien Gebhard y sus colaboradores creen estar enunciando algo científico—o sea un diagnóstico psiquiátrico—en realidad están comunicando sencillamente que las mujeres que declararon (o creyeron) que el aborto no las había traumatizado psicológicamente fueron más que las que dijeron (y/o creyeron) que sí las había traumatizado. El único—pero decisivo—error es aquí que estos autores no advertieron a qué universo de discurso pertenecían sus datos.³ Dicen por supuesto que pertenecían al campo de la psicología, pero no reducidos, por ninguna cuantía de diferenciación, a ecuaciones diferenciales (Donnan, 1956-57).

¹ Deliberadamente me olvidé aquí de las especulaciones puramente filosóficas, que influyeron en la ciencia antigua menos de lo que algunos historiadores de la filosofía parecen creer. Véase el Apéndice a este capítulo.

² Un hecho puede interpretarse debidamente sólo después de asignarlo al universo de discurso a que realmente pertenece y fuera del cual no tiene significado.

quiatria, mientras que en realidad pertenecen al de la investigación de la opinión, sencillamente porque ninguno de sus sujetos era capaz de hacer un autodiagnóstico psiquiátrico válido. De ahí, en este respecto, que los autores no registraran nada relacionado, ni remotamente con el tema que querían investigar. En lugar de ello hicieron una contribución importante —si bien no intencional— a la solución de un problema que no trataban de analizar: el problema no psiquiátrico, sociocultural, del "folklore" del aborígeneo en la sociedad norteamericana. Estas críticas serían válidas aun que más adelante un equipo de psiquiatras volviera a examinar a aquellas mujeres y descubriera que efectivamente, sus autodiagnósticos fueron atinados. Porque incluso en ese caso, sólo las apreciaciones de los psiquiatras serían datos auténticamente psiquiátricos, y las declaraciones (ahora debidamente confirmadas) de las mujeres respondientes seguirían siendo "opinión" y aun "folklore".

Defectos metodológicos semejantes afean también otros estudios de la ciencia del comportamiento conformados de acuerdo con las técnicas de las ciencias exactas, pero no inspirados en métodos científicos básicos. Además, muchos de esos estudios emplean procedimientos de apariencia científica —o mejor dicho propios de la física— no porque los científicos del comportamiento traten de hacer ver que su disciplina es tan "científica" como la física. La cuantificación, en pos del prestigio de lo incuantificables, es un acto comparable al empeño de Leibniz en demostrar matemáticamente la existencia de Dios.

Es lícito esperar que venga un tiempo en que los datos de la ciencia del comportamiento sean exactos y cuantificables. No es posible apresurar el advenimiento de ese tiempo feliz desdénando la construcción de un esquema conceptual apropiado que haga de base para la edificación, en lugar de empezar desde un techo impropio, mal ajustado y prestado, para abajo. Una ciencia científica del comportamiento sólo puede crearse recurriendo sistemáticamente a un método científico generalizado y a una epistemología generalizada, no específica de una disciplina. No puede construirse imitando las técnicas estrictamente vinculadas a la materia objeto de estudio de ciencias que tratan de fenómenos no cronohistóricos, descriptibles en función de unas pocas variables fáciles de cuantificar.

La ciencia científica del comportamiento debe empezar por el significado ni pertinencia científicos. Por ejemplo, el suicidio de Cleomenes I, rey de Esparta, no es convincente mientras el relato de Herodoto (6.75) se considera un documento histórico, pero es completamente convincente en cuanto se lo considera como el informe de un caso psiquiátrico (Devereux y Forrest, 1957). Problemas a fines se examinan también en otro estudio (Devereux, 1953).

escrutinio de la matriz completa de significados en que todos sus datos pertinentes están incluidos (Devereux, 1957a), y con una especificación de los medios con que el investigador puede acceder a, o si no educir, cuantos significados sean posibles.

El segundo paso consiste en el estudio del interés afectivo personal del científico del comportamiento por su material y las elaboraciones de la realidad que actúan esas reacciones de "contingencia", ya que el mayor obstáculo a la creación de una ciencia científica del comportamiento es el interés emocional, ineludiblemente aplicado, del investigador por su material, que en definitiva es el mismo y que por eso suscita angustias inevitables. El tercer paso consiste en el análisis de la naturaleza y el lugar del delimitado entre sujeto y observador.

El paso cuarto y último (provisionalmente) que puede darse, dado el estado actual de nuestro conocimiento, es la aceptación y el aprovechamiento de la subjetividad del observador y del hecho de que su presencia influye en "trastornos" el comportamiento de un electrón. El científico del comportamiento debe saber reconocer que nunca observa el hecho comportamental que "se hubiera producido" en su ausencia ni oye una comunicación idéntica a la que el mismo narrador hubiera hecho a otra persona. Por fortuna, los llamados "trastornos" o "perturbaciones" creados por la existencia y las actividades del observador debidamente aprovechados, son las piedras angulares de una verdadera ciencia del comportamiento y no —como suele creerse— contratiempos deplorables, con los que lo mejor que se puede hacer es esconderlos apresuradamente debajo de la alfombra.

Aunque la aclaración de estos problemas no traiga un milenio de perfección científica, el escrutinio con espíritu crítico de las dificultades inherentes a una ciencia determinada casi siempre revela el campo de su pertinencia y definen su naturaleza y con ella, los datos clave de esa disciplina.

Dicho con más sencillez, siempre ayuda descubrir exactamente qué es lo que uno está haciendo en realidad.

Las opiniones presentadas en esta obra son, en principio, aplicables a todas las ciencias del comportamiento. Yo he sacado la mayoría de mis ejemplos del campo de la sinopsicología, en parte porque sus datos son los más múltiples y determinados y en parte porque conozco mejor ese campo. Además, esta facultad de

Los tribunales de derecho, como los científicos de la conducta, prefieren ignorar el problema.

escoger es legítima. Lo que en materia de método se aplica a una ciencia cuyos datos sean un tejido complejo de variables biológicas, psicológicas, económicas, históricas, sociales y culturales y cuyo marco de referencia abarca al individuo y al grupo, necesariamente se aplica también a todas y cada una de las ciencias segmentales del comportamiento, consideradas "casos límite".

APÉNDICE

LA RELACIÓN ENTRE LÓGICA Y LABOR CIENTÍFICA EN LA GRECIA ANTIGUA

Carta de W. K. C. Guthrie, Laurence Professor de Filosofía Antigua en la Universidad de Cambridge, fechada el 10 de mayo de 1965.

Quisiera darle la declaración *ex cathedra* que me pide, pero la relación entre lógica y labor científica en la Grecia antigua es complicada y ha suscitado muchas controversias, sobre las cuales no puedo decir que tenga yo mismo una opinión firme.

Si verdaderamente quiere usted que sea breve y dogmático, me siento inclinado a decir que sí, que tiene usted razón: en sus comienzos al menos el método científico fue heurístico y no efectivamente influido por la lógica pura; pero sin duda esta afirmación necesita muchos matices. Antes de Aristóteles, su fundador, podía decirse con toda justificación que no había lógica pura; su formulación del silogismo fue realmente el comienzo de ella. De ahí que no se pueda hablar de su influencia en los pioneros del pensamiento científico o de la filosofía natural, o sea los presocráticos. Pero aun así, uno vacila: ¿era la insistencia de Parménides en la unidad e inmovilidad de lo real cosa de "lógica pura"? De un modo muy elemental, lo era. Sin duda no se basaba en la observación, pero causó una impresión tan profunda que influyó en toda la orientación del pensamiento científico después de su época.

Aristóteles es a todas luces la figura clave para su propósito, y si, como se acostumbra, uno considera que el origen de la lógica pura o formal está en su formulación del silogismo, es evidente que no tuvo mucho efecto en su labor biológica. Si por otra parte

¹ Se ha visto incluso (Meyer, 1935) que la falacia puede considerarse un cruce fuera del alcance de este libro, cuyo análisis queda

incluimos los principios de la clasificación, las nociones de *genos* y *eidos*, inmediatamente aparece una relación estrecha, pero de ninguna manera simple. Se ha dicho últimamente, por ejemplo, que la teoría y práctica aristotélicas de clasificación zoológica en el cuerpo principal de su obra biológica pueden distinguirse en claridad, tanto del método de división recomendado en *Analytics posterior*, 213 y *Metaphysics*, 2.12, por ejemplo, como del método deductivo descrito en *La Política*, 4.4 (G. E. R. Lloyd: "The development of Aristotle's theory of the classification of animals", *Phyllosophical Studies* 6: 59-81, 1961.) J. M. De Biordi: *Logique et méthode chez Aristote*, Paris, 1939, es quizá el único libro moderno que hace un estudio comparado de la teoría y la práctica de Aristóteles, y de Marjorie Grene: *Portrait of Aristotle*, Chicago, 1963. La autora es también bióloga y hace resaltar la relevancia de los intereses biológicos de Aristóteles en su filosofía. Contiene bastante de su concepto del método científico y su aplicación a su labor científica.

Quando llegamos a autores como Arquímedes y Eratóstenes, ya no estamos, naturalmente, en los comienzos del método científico, y está operando la lógica de Aristóteles y los estoicos. Preguntar usted si la obra experimental y matemática de un Arquímedes o un Eratóstenes acusa la influencia de la lógica pura, y yo le digo que con toda seguridad hay una enorme diferencia entre la obra experimental y la matemática. En Grece, como ahora, las matemáticas y la lógica estaban muy íntimamente ligadas (y tal vez sea pertinente al respecto un breve artículo de K. Berka: "Aristoteles und die axiomatische Methode", *Das Altertum*, 9: 200-5, 1963), pero es cuando menos discutible el que en la ciencia influyeron poco los adelantos de la lógica pura.

Puesto que esta cuestión sólo recibirá una breve mención de pasada, creo que lo que usted dice (en su carta) acerca del método científico, en sus comienzos es más o menos correcto. A partir de Aristóteles, la posición es más compleja y hay que distinguir de lo menos entre las ciencias exactas y las naturales o experimentales; pero incluso con Aristóteles, probablemente podríamos decir que su método biológico no fue determinado por sus adelantos en la lógica pura. (Algunos podrían decir que fue al revés.) Esta cuestión es difícil de resolver con una respuesta breve.

DEFENSAS PROFESIONALES

Todo científico del comportamiento tiene a su disposición ciertos marcos de referencia, métodos y procedimientos que —entre paréntesis— también reducen la ansiedad ocasionada por sus datos y le permiten así funcionar debidamente. Por desgracia, precisamente porque reducen la ansiedad, esos expedientes suelen transformarse de modo sistemático en verdaderas reacciones contratransferenciales que conducen a un setting que autoconstruye que se disfraza de ciencia. Es condición sine qua non para su aprovechamiento genuinamente científico y sublimatorio el entendimiento cabal del empleo neurótico que puede darse a esos expedientes.

Muchas defensas profesionales son simplemente variedades de la defensa de aislamiento, que "descontamina" el material anxiógeno reprimiendo o negando su contenido afectivo y humano así como su pertinencia personal. El siguiente ejemplo clínico aclarará la índole y la función del mecanismo de aislamiento.

Caso 10: Un paciente inicialmente sifilítico había logrado disminuir su ansiedad tan eficazmente de las enfermedades venéreas que era capaz de cohabitar incluso con prostitutas de los tugurios sin sentir ninguna angustia. Pero en cambio tenía un miedo pánico a otras enfermedades, como la rabia, la meningitis cerebros-pinal y la poliomielitis, que, como la sifilis terciaria, atacan al sistema nervioso, y se desinfectaba meticulosamente las manos aun después de haber tocado un periódico o un libro donde estuviera la palabra "poliomielitis".

El psicoanalista se preparó sistemáticamente a trabajar con el material causante de ansiedad sometiénolo a un análisis didáctico, para solucionar sus propios problemas. Esto le suele permitir aguantar sin angustia indebida el bombardeo a que el material anxiógeno producido por los pacientes lo somete y esudarlo sin sentir que desfigurarlo, de modo que pueda controlar sus propias ansiedades. Además, si su análisis ha sido verdaderamente bueno, tiene el insight suficiente para enviar a un paciente cuyos problemas no puede manejar con la debida calma a otro analista que tal vez esté en condiciones de tolerar mejor que él el material

anxiógeno producido por ese paciente. Las reglas del procedimiento psicoanalítico ayudan también al analista a ser objetivo, como su autodefinición ("soy un analista") y la definición de su situación ("esto es un análisis") (Dévereux, 1956a). Por ejemplo, es capaz de reaccionar objetivamente a las observaciones insultantes o amenazadoras de un paciente, porque esas definiciones de su papel y de la situación le ayudan a comprender que las observaciones del paciente son manifestaciones de un enojo transferencial y en realidad van dirigidas contra una persona que desempeñó un papel de importancia principal en la infancia del paciente. ✓

1. Preferencia vicaria: El impacto traumático de un suceso, potencialmente anxiógeno, se reduce de un modo apreciable si uno está preparado; ya como dice William James, uno tiene "conocimiento" anticipado de la situación está en condiciones de "trabajar conocimiento" con ella. Por ejemplo, el estudio de la antropología en la universidad fortalece al antropólogo para el impacto de la experiencia real sobre el terreno de los hechos, aunque no puede, naturalmente, hacerlo del todo inmune al impacto del rito de subincisión australiano. Sencillamente lo ayuda a retrasar (postponer) la experiencia consciente de la angustia hasta que haya fotografiado y registrado ese ritual. Además aunque el caso #2 muestra que las experiencias de campo traumáticas acaban por producir una reacción de ansiedad diferida de cierta intensidad, lo que importa científicamente es que sin ese "conocimiento anticipado" de esas experiencias el antropólogo no hubiera podido observar y describir tales prácticas con exactitud.

Incluso algunos neuróticos saben que es posible reducir la intensidad de un ataque de ansiedad "anticipándolo", aunque en una forma algo extrema:

Caso 11: Un joven con una fobia aguda, contra los puentes decía: "Cuando sé que tengo que atravesar un puente, deliberadamente me hago tener un ataque de ansiedad antes, porque sé que así disminuye la angustia que siento al pasar el puente. Por eso no temo el sentirme angustiado antes; la ausencia de ansiedad en el momento me hace aún más ansioso de lo que estaría normalmente al cruzar el puente."

La posición profesional y la defensa de actividad se combinan en una definición de sí mismo ego-sintónica y culturalmente sancionada ("soy un antropólogo") y una definición similarmente sancionada de la situación ("esto es trabajo de campo"). Mientras que la protección proporcionada por la posición cientí-

fica es sólo temporal, y con frecuencia dura tan sólo mientras uno está efectivamente dedicado a la labor científica, la actividad científica en sí puede atenuar ulteriormente la experiencia consciente de la ansiedad al permitir una abstracción parcial de las tensiones por la actividad. De todos modos, en cuanto esa era, la ansiedad no experimentada conscientemente antes puede manifestarse con una intensidad abrumadora, característica del "retorno de lo reprimido". Además, si la índole de la actividad científica es tal que produce tensiones por sí, la ansiedad puede acabar por ser tan indomeñable que requiera una intervención psicoterapéutica (caso 257). Queda descrito (caso 39) el modo en que mis propias actividades de comentarista me permitieron detener la inquietud mientras pasaba una película causante de ansiedad. Sobre el terreno pueden ocurrir también experiencias semejantes.

Caso 12: Dos antropólogos no sintieron "ninguna ansiedad" mientras observaban, grababan y fotografiaban una circunsción femenina africana sobre el terreno. Pero cuando vieron después la película Karamoja —que mostraba los ritos, idénticos, de circunsción femenina de otra tribu— sintieron una "angustia tremenda", sencillamente porque su papel de espectador pasivo no les daba esta vez ocasión de abstraccionar la angustia por medio de la actividad.

El científico del comportamiento que se encuentra frente a materia traumática no tarda en aprender a utilizar la posición profesional como procedimiento reductor de la ansiedad, sobre todo el lado que puede —dentro de ciertos límites— incluso entregarse a actividades que normalmente le causarían intensos sentimientos de culpabilidad. Y así Malinowski (1932) pudo estudiar la vida sexual de los isleños de las Trobriand y yo la de los mohave sin sentirnos molestados y sin que nuestras investigaciones nos estimularan indudablemente. Sospecho empero que los sujetos que se presian voluntarios a ciertos tipos de investigación acerca del coito humano (casos 121, 122) simplemente aprovechan la definición científica de la situación para fines de *acting out* neurótico y aun perverso (exhibicionista).

Sea como quiera, la creciente aceptación de la ciencia de la conducta como una forma legítima de "hacer ciencia" puede explicar en parte por qué las monografías antropológicas contemporáneas describen de un modo muy franco material "escabroso" que en las monografías anteriores se omitía decididamente, se describía en la ínfima, o se mencionaba sólo con eufemismos y de un modo que hacía constar el horror y el disgusto del observador. (En algunas expo-

siciones antiguas de las prácticas de subincisión australianas se las llamaba simplemente "ese terrible rito".)

Por desgracia, esta posición profesional culturalmente de ahora ha llevado el péndulo tan lejos hasta el otro extremo que algunos antropólogos consideran "meras costumbres", no sujetas a estimación ética, algunas prácticas crueles e irracionales. Ambos extremos representan reacciones contrasféricas neuróticas; ninguno de los dos es verdaderamente objetivo ni científico. Pero sea como quiera, el caso 59 muestra cómo la posición científica nos puede permitir hacer sobre el terreno lo que ni en sueños haríamos en otro contexto y además, hacerlo de modo sublimaricio, o sea de modo que el resultado sea objetivamente menos malo, y subjetivamente menos agotador, de lo que hubiera sido si uno se hubiera abstenido decididamente de obrar. Ciertamente, es a menudo posible hacer arreglos que le permitan a uno estudiar una práctica desagradable sin perjudicar a ningún ser viviente y sin exponerse sin necesidad a una visión que al despertar, nuestras angustias reduce la precisión de nuestras observaciones.

Caso 73: Los sedang castran a los verracos con una astilla aguda de bambú y no con un cuchillo de hierro, para que los espíritus no crean que se los van a sacrificar a ellos. Como los sedang son duros en su trato a los animales y a veces castran a los perros sencillamente por "divertirse", no tenía yo ningún deseo de presenciar la torpe operación de captar a un joven verraco. Entendí, la siguiente vez que me mataron a un verraco los pedí que castraran el *caddar* exactamente igual que hacían con los verracos vivos. Esto me permitió hacer observaciones más precisas que si me hubiera visto obligado a tomar notas mientras me distraía —y allí, naturalmente, que este subterfugio me impedía observar también las reacciones psicológicamente importantes del capador. Tomé esto en cuenta eligiendo para la tarea a un excelente mimico, que estoy seguro manifestó exactamente las mismas emociones de un sedang que castrara un verraco vivo.

Las posiciones metodológicas y los requisitos técnicos que cuando se emplean como es debido, son lógicamente imprescindibles y científicamente fructuosas, pueden emplearse en forma inconsistente con todo como defensas aisladoras que desfiguran nuestra percepción de la realidad y estorban la investigación de varios modos.

2-2. El relativismo ingenuo, cultural y ético —una concepción de la humanidad como "museo de las costumbres"—, reconstruye la exis-

tencia de seres humanos, pero en nombre de la objetividad "científica", se niega a aplicar las consideraciones éticas ordinarias a su conducta. Además, esta posición "metodológica" esquivaba el importante problema de la relevancia del ethos de una cultura para otra, así como el importante problema cultural de la ética en general.

Caso 44: Un antropólogo psicológicamente experto me dijo haber presenciado el entierro de una persona viva que había "perdido su alma" y por eso era considerada técnicamente muerta (cf. caso 275). Cuando le pregunté si no había tratado de impedirlo me replicó altanero: "En mi calidad de antropólogo, mi obligación no es acabar con las costumbres de los indígenas, sino estudiarlas."

La negación compulsiva de la legitimidad de los juicios éticos puede conducir a negarse a hacer diagnósticos científicos aun cuando el grado de patología de una sociedad dada pueda evaluarse en función de su fracaso en alcanzar la meta que declara y de su propensión a destruirse a sí misma por la adhesión a metas disfuncionales. El mejor ejemplo es la tribu de los tonkawa con su disposición patológica autodestructiva a exponerse a que la "extermine" sus vecinos indignados en lugar de renunciar a su canibalismo, económicamente innecesario (Linton, 1947; Devereux, 1955b).

El relativismo cultural trata así de reducir la ansiedad considerando los datos culturales en un vacío humano. Aunque científicamente estéril, este subterfugio es efectivo, puesto que reacciona con menos ansiedad a la canibalización colital de su macho (con el que no podemos identificarnos) por la mantis hembra de lo que haríamos a una costumbre semejante entre los humanos. De modo análogo, podemos reducir artificialmente nuestras angustias viendo torturar a prisioneros tan sólo como "costumbre" u eligiendo con ello implícitamente que esas prácticas tengan alguna relación alguna con seres de carne y hueso, con los que tendríamos que identificarnos. Estas negaciones implícitas de toda semejanza entre nosotros y otros también se practican en la vida corriente, por ejemplo, en los intentos de justificar la esclavitud, que, "en realidad", sólo afecta a unos "semianimales". Este aumento intencional de la distancia social entre el antropólogo y los indígenas, que estudia le permite olvidar sus propias angustias, estudiando sus costumbres como si la cultura no afectara a las vidas humanas, aun cuando que no es necesario decir que esas angustias denegadas son desplazadas sobre otros temas. Muchos antropólogos son hostiles a los que estudian las costumbres en relación con el hombre, sobre todo porque este enfoque más amplio amenaza con reintroducir el elemento humano (psicológico) causante de ansiedad en su mundo.

cuidadosamente esterilizado de afectos, en que sólo hay costumbres e instituciones "puras".

El modo de ver "relativista" de que una práctica cruel es sólo una costumbre éticamente neutra puede irrazonablemente impedir la objetividad verdadera, sencillamente porque la definición relativista de la costumbre no suelen compartirla los que de hecho la practican, que pueden considerarla moralmente errónea (Lévi-Strauss, 1955). El negarse a advertir cuán viciosa es en sí una costumbre puede por eso no dejarnos ver la evaluación acostumbrada de esa práctica por el primitivo.

Caso 45: Si uno considerara la masturbación premarital de los sedang, su homosexualidad, bestialidad y sus perversiones heterosexuales como "costumbres nada más", nunca descubriría que los mismos sedang consideran esas prácticas objetables y sólo se entregan a ellas porque sus dioses, a los que explícitamente declaran éticamente corruptos y malos, preferieren prohibir las relaciones heterosexuales premaritales, que los sedang consideran en cambio moralmente más justas que las perversiones. A su vez, este descubrimiento es indispensable para comprender la hostilidad de los sedang para con sus dioses (Devereux, 1947c).

La negativa del trabajador de campo a tomar nota de que los mismos primitivos desaprueban algunas de sus prácticas tradicionales puede así impedirles predecir o comprender el cambio cultural que representa una reacción a una práctica tradicional pero mal vista.

Caso 46: Cuando, en el pasado siglo iban a sacrificar a un cautivo al lucero de la mañana, un gran guerrero pawnee, reaccionó y liberó a la futura víctima declarando que ya no se sacrificarían más cautivos. La tribu lanzó un suspiro colectivo de alivio y decidió abolir el rito (Linton, 1923).

Caso 47: Muy contra su voluntad, el pueblo de Temese se veía obligado cada año a consagrar la virgen más hermosa al malvado héroe muerto Polites. Pero el boxeador y paracatista olímpico Euthymos, que se apiadó de una de esas jóvenes y se enamoró de ella, desafió y venció a Polites, quien se lanzó (es decir, su estatua) al mar, liberando así para siempre a Temese de la obligación de hacer ese sacrificio (Pausanias, 6.6.2; Estrabón, 6.1.5, p. 255; Aclimano, *Varia Historia*, 8.18; Eustatio: *De la Odisea*, 1409.13).

Caso 48: El sentimiento subjetivo que tienen los sedang de que comerse el hígado humano es moralmente repulsivo hito que se reemplazara esa práctica, al principio con un gesto puramente simbólico —tocar un hígado humano con los labios— y después por el consumo del hígado de un animal sacrificado.

Caso 49. El sacrificio humano también se hizo obsoleto entre los sedang nada más porque les parecía moralmente errado. Al principio limitaron los sacrificios humanos a las ocasiones en que una crisis muy grave — como una epidemia — hacía parecer necesario el recurso a esa antigua costumbre. Después, hacían como si sacrificaran un ser humano, pero en realidad sólo le punzaban el sobaco y mataban un puercito en su lugar.

Además, cuando un devoto "sacrificado" de esta manera ve mandó a sus "matadores" por daños y los ancianos le organizaron una compensación, los sedang renunciaron incluso a fingir el sacrificio de un humano. Para 1933 ya habían reemplazado las víctimas humanas por figurillas de cera y se contentaban con matar un búfalo. Incluso racionalizaron esta práctica proclamando orgulloosamente que ellos eran más ricos que los espíritus, que tenían que comer seres humanos mientras que ellos se permitían comer búfalos!

Estas dos innovaciones culturales ocurrieron antes de que los sedang fueran sometidos a una presión aculturativa o administrativa. Se produjeron, porque, a diferencia de ciertos antropólogos, los sedang no consideraban que comer carne humana o matar gente fuera "nada más una costumbre". De hecho, se alegraron cuando las presiones administrativas los obligaron a abandonar también otra costumbre onerosa.

Caso 50. Cuando los franceses prohibieron las guerras entre tribus y obligaron a los sedang a reemplazar las incursiones rituales por simulacros, un sedang me dijo: "¡Es mucho mejor así! Ahora podemos gozar del festín de la victoria y emborracharnos a fondo sin tener que jugarlos primero la pelleja peleando con los halang." Estos datos muestran que la actitud de "nada más una costumbre" suele hacer que el investigador, con pulverizante "objetivo" no sea nada objetivo acerca de las costumbres que estudia.

2b. El eliminar al individuo de los informes etnológicos de la capacidad anteriormente un procedimiento acostumbrado.
Caso 51. Linton me mencionó una vez la siguiente observación

También Temiwoets durante las guerras Médicas se vio obligado a hacer sacrificios humanos (Plutzer, Temiwoets, 13); después de su derrota final por Agrotel los cariangs volvieron a sacrificar a sus propios hijos en lugar de otros comprados (Trudner, Siulo, 20144). Un sacrificio humano comparable se practicó, según Euripides (*Higiena en Tauride*, versos 1458 ss.), en Hala Arraphidas ante la estatua de la Artemisa local, que recibía sacrificios humanos verdaderos en Taurida. Como dijo una vez con ironía Hallowell: "A los antropólogos, sencillamente, la gente no les interesa mucho."

de un colega: "Mi monografía de la tribu X está casi acabada. Todo cuanto me queda por hacer es quitarle la vida" (o sea, todo lo relacionado con personas y sucesos reales).

2c. Los esquemas conceptuales válidos y las posiciones metodológicas pueden emplearse también, primordialmente, para la descomposición selectiva del material anxiógeno.

La etnología suele operar como si la gente no existiera en realidad. Claro está que este procedimiento preliminar es igualmente legítimo para estudiar la cultura aislada — sea *in vitro*, sea como una "variable analítica" — con tal que los resultados de este estudio segmentario del comportamiento humano se correlacionen en definitiva con los descubrimientos y formulaciones psicológicos.

¶ Pero si se piensa que la etnología ha de dar soluciones finales y compendiosas y si hay el temor — completamente injustificado — de una conjura para reducir lo sociocultural a lo psicológico (Kroeber, 1918a), entonces la posición etnológica o la psicología profesional temporal con un fin determinado.

2d. El enfoque atomístico debe distinguirse claramente del empleo de listas de rasgos y cuestionarios como artificios metodológicos que garanticen que el trabajador de campo no olvidará "ningún accidente o intención" el material que — a causa de sus ansias de preferiría no investigar. Hay que distinguir también de las buenas estudios de distribución que — como el de Kroeber (1919, 1932, 1935), análisis de zonas culturales americanas y tipos de personalidad por áreas — no ignoran la configuración de rasgos y su pertinencia humana, así como de las investigaciones acerca de la amplitud y variedad de rasgos culturales, que en definitiva tratan de arrojar luz sobre el funcionamiento total del hombre en cultura (Devereux, 1932a). La ciencia atomística de la conducta representa una defensa contra la ansiedad sólo si, después de arrancar un rasgo de su contexto, no lo reintegra finalmente a su matriz psicológica y cultural.

¶ Yo mismo eboet (Devereux, 1940), un esquema conceptual de sociedad que eliminaba de momento al individuo, pero sólo para esclarecer la naturaleza del hecho sociológico. Después (1945, 1957a, 1961b) distinguí la interdependencia y la relación de complementariedad existentes entre el entendimiento sociológico y el psicológico de un acto dado, pero insistí (1956, diciembre de 1961b) en que eso era sólo un paso preliminar hacia la formulación de un marco de referencia unificado, en que se combinarían esos *intights* segmentarios teóricamente distintos en un esquema único. En ninguna parte insistí que esos *intights* no pudieran ni debieran unificarse al final.

mientras que algunos otros modelos, como el de Erikson (1930) siguen —por lo menos para mí— siendo esquemas exangües, aunque quizá deslumbrantes. Es instructivo comparar en este contexto la vida de los modelos altamente complejos, construidos por Röhrl (1930), Mendel (1947a, b, c), Lévi-Strauss (1949, 1955, 1962a, b, 1964) o La Barre (1954), con el brillo helado del modelo formalista que traza Ersson (1943) de la cultura y personalidad yurok o —en grado algo menor— también con algunos (pero todos) aspectos del modelo de la personalidad étnica creado por Kardiner (1939, 1945). Ciertamente, incluso aquellos que no aceptan los modelos de Röhrl, Mendel, Lévi-Strauss o La Barre sentirán alguna que otra vez que esos modelos en apariencia abstratos evocan, con una penetración casi sobrecogedora, un amigo, un paciente o un informante primitivo. En cambio, un modelo puramente intelectualista tan sólo —deslumbrante— al intelecto. Es tan "fino" como una disquisición medieval acerca de cuántos ángeles caben en la punta de una aguja; sólo nos espanta por su ingenio pero no nos sigue hacer que los ángeles ni la aguja —ni tampoco el autor del tratado— nos resulten reales, puesto que no dejan pensando "¿y eso qué?".

Un mismo procedimiento teórico genuinamente importante, si se emplea a manera de defensa contra la objetividad, puede convertirse en una de las cosas advertidas —por ser narcisísticamente más humillantes— causas de distorsión en la ciencia de la conducta: es un tipo particular de construcción de teoría en dos pisos.

1. La primera fase consiste en formular una teoría que explique debidamente aquella parte de los hechos que menos ansiedad causan. Esta teoría segmentaria suele entonces servir para desalentar la investigación en la otra porción de los hechos: la que más ansiedad causa.

2. En la segunda fase, esta teoría segmentaria se elabora sistemáticamente para crear la ilusión de que es completa, con lo que desalientan intentos ulteriores de encerrar los aspectos inquiridos de los hechos que uno profeta haber explicado.

Aunque el valor inherente —si bien transitorio— de las teorías que explican un aspecto de un proceso sólo está por encima de sus limitaciones, lo que importa aquí es que la teoría parcial (correcta) se materializa en los trabajos publicados. Se construyen de modelos —para no mencionar la de Freud— puede compararse con provecho con modelos puramente intelectuales en la psicología psicoanalítica del Yo, como los de Rapoport (1951a, b).

aplica entonces para impedir la formulación de una teoría verdaderamente completa.

Caso 54: Hace unos años me interesaba yo más en los padres de Edipo que en Edipo mismo. Releyendo los textos griegos sobre Edipo como si no trataran primordialmente de Edipo sino de Layo y Yocasta descubrí que estos mitos consideraban explícitamente la conducta "edípica" de Edipo consecuencia directa e inevitable del (y reacción al) comportamiento de Layo y Yocasta. Esto implica que el complejo de Edipo podría con toda legitimidad llamarse complejo "contralayo" o "contrayocastas" y que el complejo de Layo y el de Yocasta se llaman indebidamente contradidípicos. Por cierto que esta denominación da a entender erróneamente que el complejo del hijo Edipo provoca los complejos "contradidípicos" de sus padres. Inecesario es decir que mis observaciones no ponían en duda la existencia del complejo de Edipo; en realidad lo confirmaban considerablemente. Indicaban tan sólo que el complejo del hijo Edipo es ante todo la reacción provocada por los impulsos incestuosos y/o asesinos de los padres y por eso es, hablando en puridad, un complejo de contralayo o contrayocasta.

Aunque este trabajo manejaba hechos demostrables, se me aconsejó en forma privada que no lo publicara porque podría perjudicar a mi reputación de freudiano clásico, entonces firmemente asentada. En realidad, el órgano oficial de la Asociación Psicoanalítica Internacional aceptó el artículo casi a vuelta de correo (Devereux, 1953b).

Caso 55: Otras investigaciones de los aspectos científicamente escolomizados de otras complejas relaciones de la infancia me vieron a redactar un artículo sobre el carácter derivado de los impulsos caníbales en la primera infancia, en que asentaba los siguientes puntos, copiosamente documentados:

1. Nuestra información sobre la vida psíquica del infante es en gran parte inferencial y se compone de reconstrucciones más o menos válidas.

2. El lactante no es todavía capaz de distinguir entre la carne humana y la no humana; por eso no puede, en sentido psicológico, decirse que tiene impulsos caníbales. Se reconoce que inunde el pecho, pero con el mismo placer muere el chupete. Por eso puede

1. A la milha conclusión llegó implícitamente hace tiempo el latente Buhler (1891), quien hizo ver que los medios por los que según una tradición típicamente a los infantes correspondían el tipo de castigo que se aplica del siguiente pariente de Edipo.

Incluso los estudios acerca de la índole de la cultura, de los procesos culturales y de difusión pueden ser viciados por la negativa a tomar en cuenta la psicología.

Caso 52: Norbeck (1955) trató de explicar la presencia del rezago de hombres sin ano en dos regiones muy separadas una de otra mediante una hipótesis extravagante de que la habían introducido en las Filipinas esclavos llevados de América del Sur. Si Norbeck hubiera tomado en cuenta también la creencia chaga (Raum, 1939) de que los hombres adultos no tienen ano y las fuentes griegas y romanas como Luciano (*Verba Historica*, 1.23), Plinio (*Naturalis Historia*, 7.25) y Aulo Gelio (*Noctes Atticae*, 9.4.10) hubiera comprendido que no era posible la solución del problema exclusivamente en términos de difusión, sencillamente porque la negación del ano es una fantasía humana que puede manifestarse en muchos niveles: como cuento en América del Sur y las Filipinas, etc., como creencia impuesta socialmente entre los chagas y como fantasía neurótica en los pacientes occidentales (Keiser, 1954; Devereux, 1954b). Además, el mito sudamericano también puede explicarse estructuralmente (Lévi-Strauss, 1966).

Caso 53: Si uno desea estudiar la idea del pene luxado etnológicamente tiene que considerar como un solo universo de discurso una práctica tupari, una broma mohave, un mito zoroástrico, una mentira esquimal, una neurosis koro indonesia y del sur de China, un cuento de *Les Cent Nouvelles Nouvelles*, una anécdota semi-pornográfica de la monarquía francesa, la acción de un psicótico alemán, el sueño de una neurótica norteamericana, las paratestias de un bostoniano *borderline*, la fantasía de un obsesivo del medio oeste, etc. (Devereux, 1954a, 1957a). Sólo un mapa de distribución que consigne todas esas manifestaciones de una misma idea puede tener sentido antropológicamente... y la preparación de ese mapa requiere que se tenga conocimiento del hombre en toda su realidad corpórea y psíquica.

Caso 54: La psicología de rasgos, que malemplea los tests como fajas cortafuegos emocionales es la contraparte psicológica de la ciencia social atomística. La capacidad que tienen los tests de proporcionar un entendimiento intelectual y al mismo tiempo de incrementar, no reducir, nuestro sentido de la realidad humana y psicológica del sujeto sometido a test se prueba, en el campo clínico, por los informes de test de los mejores psicólogos y en el campo de la cultura y la personalidad por las investigaciones de trabajadores inspirados, como Hallowell (1955): En manos de las personas que no temen a las realidades psíquicas, los tests no son escapatorias sino sendas que conducen hacia el entendimiento de los seres vivos.

Por desgracia, hay también algunos "expertos" que (inconscientemente) se sirven de los tests para esterilizar su material.

La monografía de Lessa y Spiegelman (1954) sobre la personalidad utiliana, que se basa primordialmente en datos de tests, así como dos libros basados en parte en los tests realizados con navajos (Kluckhohn y Leighton, 1946; Leighton y Kluckhohn, 1947), no lo gran dar vida, al menos para un lector (Devereux, 1948a, f) a los individuos y el grupo que describen, mientras que Lowie (1935) —que no profesaba el ser psicólogo ni se servía de tests— anima tanto a la tribu *crow* como a sus miembros. Incluso una obra tan teórica como el excelente estudio de Llewellyn y Hoebel (1941) sobre las leyes cheyennes, escrito por dos eruditos no especializados en cultura y personalidad, y que persiguen fines diferentes, da vida a los cheyennes, a pesar de diagnósticos psiquiátricos incorrectos a veces.

De hecho, algunos antropólogos básicamente antipsicológicos parecen haberse convertido en "especialistas" de cultura y personalidad como una solución de transacción ("acatamiento fingido") y haber publicado labor estadística de tests principalmente para tener la "autoridad" necesaria para combatir todo enfoque psicológico verdaderamente profundo del hombre. Algunos trabajos de Lessa (1956), Orlansky (1949) y Sewell (1952) son buenos ejemplos de esta maniobra. La Barre (1958) cita otros datos al respecto.

Caso 55: La construcción intelectualista de modelos de personalidad es la contraparte psicológica de la teorización cultural. La construcción de modelos es una práctica normal en cualquier ciencia que ha llegado al punto en que la teorización se hace necesaria y factible. Pero debe comprenderse que no es posible construir modelos abstractos de la personalidad sin "descontaminar" primero nuestro material. Los modelos buenos, por muy abstractos que sean —y algunos de ellos son muy abstractos, ciertamente!— pertenecen a seres humanos y no a "preparaciones" de laboratorio. De ahí que le "prendan el foco" tanto al lector ordinario como al clínico:

* Los clínicos dinámicos suelen quejarse de que los informes de tests de los psicólogos clínicos de segunda categoría son esquemas vacíos de sentido y repletos de una jerga que hace el papel de *injuigit veritatis* y empatía clínica creadora.

* Una mujer cheyenne anormalmente mandona y agresiva, a quien Llewellyn y Hoebel califican de "psicópata", era casi con toda seguridad un caso grave de carácter neurótico.

* El doctor Leewenstén me dice que siempre que él y los doctores Hartrmann y Kiri trabajaban en algún artículo de teoría comprobaban toda construcción teórica citando en el curso de sus disquisiciones pruebas clínicas apropiadas, aunque por razones de comedia en la exposición a veces no mencionan etc.

decirse que tiene impulsos de morder y devorar, pero no que los tenga canibales.

3. La creencia de que los infantes tienen impulsos específicamente canibales se debe entonces a la *proyección* sobre el niño de los impulsos canibales de sus padres, o bien a la *proyección regresada*, sobre la primera infancia de los impulsos canibales de niños algo mayores, capaces de distinguir entre la carne humana y la no humana y cuyos impulsos canibales representan una respuesta a los impulsos canibales de sus padres.

4. La zoología, la antropología y la historia por igual, demuestran que los animales adultos y los seres humanos devoran a sus pequeños, mientras que éstos nunca devoran a sus padres; y aun después de llegar a la madurez, es muy difícil que los devoren. Incluso la *manducación* del cadáver del padre o la madre muertos o moribundos (Herodoto, 3.99) es excepcional, y en cambio está muy difundido el matar a los hijos, tanto para comerse los como por ritual (Devereux, 1966d).

Estos hechos manifiestos y estas conclusiones casi-ingenualmente obvias han provocado reacciones extraordinarias.

A. Reacciones de un colega a quien enseñé el original en privado:

1. Mencionaba yo las creencias hawaianas acerca del inconsciente deseo de comer de las mujeres encinta y citaba la *palabra* hawaiana que designa esas ansias. A pesar de eso, me preguntó si esas creencias las tenían unos cuantos hawaianos solamente o la nación entera.

2. Me pidió que citara mis fuentes para algunos de mis declaraciones relativas a los mohaves, aunque era claro que yo estaba citando datos recogidos por mí sobre el terreno.

3. Aunque era médico, se "preguntaba" si sería cierto que:

a) algunas hembras de animales se comen la placenta;

b) esa ingestión promueve el paso del equilibrio endocrino de tipo preñez al equilibrio endocrino de tipo lactancia;

c) un animal experimental, privado de ciertos elementos necesarios de su alimentación, escogerá, entre diversos tipos de alimento, el que contenga los elementos que le faltan en su dieta.

4. Me pidió que citara la autoridad para mi declaración de que algunas mujeres primitivas recurren al ejercicio extremado para abortar, aunque yo citaba mi obra, *Estudio del aborto en las sociedades primitivas* precisamente para ese punto (Devereux, 1955a).

5. Puso en duda mi afirmación de que en tiempos de hambre los australianos se comen a sus hijos, aunque yo citaba dos fuentes apoyándola.

6. Me pidió que "aclarara" cierto número de observaciones extremadamente simples, etcétera.

B. Dos opinantes. Después se publicó una versión cuidadosamente reelaborada de este artículo en una publicación periódica a la que cuyos artículos acompañan los comentarios de varios opinantes. Uno de mis tres opinantes —de modo perfectamente legítimo— examinaba el problema desde un ángulo diferente y llegaba a conclusiones que considero complementarias de aquellas a que yo había llegado. Los otros dos opinantes evitaban cuidadosamente abordar el hecho crucial de que para tener fantasmas canibales hay que saber la *diferencia* existente entre carne humana y carne de animal. También equivocaban otros dos puntos importantes, que imputaban opiniones que no había yo expuesto y se conformaban más o menos con hacer observaciones sarcásticas y gritar "herejía!" (Goodley, 1966; Ekstein, 1966).

No importa aquí lo extraño y capcioso de las reacciones, de las cuales citamos tan sólo unas pocas en los párrafos que anteceden.

Lo que importa es que esas reacciones las había provocado sin ninguna duda la naturaleza *anxiógena* de mis conclusiones, que trasgredían dos prácticas científicas importantes:

1. El tabú sobre el *estrutinio* de la conducta parental en términos psicoanalíticos.

2. La tradición de que todo se le ha de achacar al paciente. Pero ya Loewenstein (1947) protestaba porque se atribuían todas las desdichas del paciente a su "masoquismo moral".¹⁰

En mi calidad de freudiano impecablemente clásico yo sugeriría que ciertos segmentos no analizados de la teoría psicoanalítica pueden servir de resistencia a los hechos (Devereux, 1965b), y esto a pesar de la insistencia de Freud de que los hechos son antes que la teoría (Freud, 1963).¹¹ Y si el psicoanálisis se hizo ciencia fue precisamente por ser capaz de aceptar los hechos que mucha gente prefiere ignorar.

El aprovechamiento de la teoría en forma defensiva no se limita al psicoanálisis.

Caso 56: La clave de la teoría de Guthrie (1935) acerca del

¹⁰ Mi rotundo y clarísimo maestro del Instituto psicoanalítico desde me paré criticaba este tabú.

¹¹ Como observó una vez un colega verdaderamente realista: "En nuestro negocio, el cliente nunca tiene razón."

¹² No menosaba la validez de esta afirmación el hecho de que Freud, de dichosamente, la hiciera en relación con la telepatía.

aprendizaje es el principio de "una-prueba-cum-posteridad" (Vocks, 1954). Se reconoce que este sistema es consecuente consigo mismo, compendioso e irrefutable pero (y por desgracia esto raramente se cumple) sólo en el limitadísimo sentido en que lo son las reglas del ajedrez, que prevén toda contingencia. Por eso el sistema de Guthrie sólo merece ese tipo de estimación semisensible, que puede "descifrarse" por medio de una maña enojosamente sencilla. Postula Guthrie que el aprendizaje se realiza por completo en una sola prueba exitosa. El hecho de que se pongan a estos tests tanto el sentido común como la experiencia diría y que sólo puede "demostrarse" por medio de argumentos en extremo espicuos es lógicamente mucho menos inquietante que el de que contradice los datos más fundamentales de toda ciencia de la vida: todo organismo es un sistema cronológico cuyo comportamiento en el tiempo t \neq Δt no es totalmente comprensible en función de su estado en el tiempo t , sino sólo en función de toda su historia anterior (capítulo 1). Este hecho no afecta a la coherencia consigo mismo, la compendiosidad e irrefutabilidad del sistema de Guthrie, pero lo relega a esa clase de sistemas operacionales cuyo prototipo es el juego de ajedrez. Esas hazañas semijuntas a juegos y casi escolásticas gustan, como es sabido, a todo cuanto hay de infantil en la mente compulsivamente "científica". Por eso sólo toman en serio algunos el sistema de Guthrie, por ser tan reduccionista que la posición olímpica del investigador no se ve amenazada por el hecho de que la rata que corre por el laberinto es tan capaz de sentir como él. Y por fin, y esto también es importante, el sistema de Guthrie es particularmente satisfactorio para cierto tipo de investigador porque no sólo lo pone fuera del experimento sino que, por encima de todo, pasa por alto el hecho de que el observador, que debería ser también sujeto según las leyes guthrieanas que rigen el comportamiento de la rata (Guthrie, 1935, 1938), pueda conservar —evadiendo todas las cuestiones psicológicas y lógicas— su identidad humana y sus experiencias cognitivas sin las cuales, como muestran los capítulos II y III, no podrían realizarse absolutamente ningún experimento, y no digamos interpretar sus resultados.

"One trial cum posterity". Según Guthrie un solo ensayo logrado asegura el aprendizaje. [A.T.]

Es digno de nota el que los psicólogos cognitivistas, que no reducen esta cuestión, hacen muchos experimentos dolorosos y destructivos en sus laboratorios de laboratorio que los psicólogos reduccionistas de república.

Y así, independientemente de que una teología privada o un reduccionismo mecanicista sean el punto de partida de intentos de esquivar las cuestiones suscitadas por la naturaleza sui generis de los fenómenos de la vida y en especial del hombre y sus obras, la diferencia entre fenómenos vitales y fenómenos no vitales ha influido radicalmente en la historia de las ideas del hombre acerca de la biosfera y de sí mismo. Pero lo que aquí nos importa es sólo el problema concreto de la índole propia de los datos comportamentales y del marco teórico que trata de la vida como fenómenos vitales y no como otra cosa.

Un rasgo común a todas las investigaciones es que en alguna parte del experimento un evento se transforma en una percepción. Alguien —sea Zeus, Einstein, Darwin, Freud, E. R. Guthrie o, para lo que nos importa, Perico de los Palotes...— dice: "y esto per-cibo" —cosa que constituye un enunciado cognitivo (capítulo XXII). Según el modo de ver de Poincaré, de que el objeto de la ciencia es descubrir las semejanzas en las diferencias y las diferencias en las semejanzas, trato de demostrar que un modo operativamente eficaz de distinguir entre datos referentes a la vida y datos no referentes a la vida es escudriñar estos dos tipos de fenómenos en el preciso momento del proceso experimental en que alguien dice: "y esto percibo."

El empleo defensivo de la teoría puede viciarse sobremedida semejante empresa, ya que, como veremos en el capítulo XXII y en el XXIII, la teoría sólo determina el punto exacto en que el experimentador u observador percipiente declara: "y esto percibo" y por eso se ve autorizado a gritar también: "Eureka!"¹²

"Suele decirse que mientras los teóricos del aprendizaje "de respuesta" crean teorías sistemáticas, los teóricos del aprendizaje cognitivo (de "situación") tratan ante todo de refutar las opiniones de los psicólogos de la respuesta. Esta acusación es sustancialmente correcta. Es evidente que los psicólogos cognitivistas evitan la construcción de sistemas porque los psicólogos pudieran remediarlo, a la posición psicoanalítica, que los pondría al margen de casi toda la psicología académica oficial; y además detendría todo diálogo ulterior con sus colegas orientados hacia la respuesta, puesto que para que haya controversia se requiere al menos una comunidad parcial de lenguaje e ideas. Tolman adopta una posición que es prácticamente la del psicoanalista clásico (Tolman, 1949).

10 LA EVALUACION

La etapa final del proceso del currículum es la evaluación basada en la medición y en la valoración. Una vez que los objetivos están claros, se han formulado operativamente, y las experiencias y el contenido seleccionados se han integrado, organizado y puesto en práctica en clase, aún se precisa una doble operación. El carácter y la amplitud de los cambios de conducta han de ser medidos en el caso de cada alumno, y habrá que evaluar tanto el cambio como la falta del mismo. Las consecuencias de los procedimientos planificados en clase se han definido en términos de conductas observables. La evaluación nos permite comparar las conductas reales con las conductas esperadas (u objetivos), y llegar a ciertas conclusiones sobre esta comparación con vistas a la acción futura. Evidentemente, se trata de una fase vital, porque, sin la comparación cuantitativa y cualitativa de las conductas reales y de las esperadas, es imposible saber si los objetivos han sido alcanzados, y en caso de haberlo sido, en qué medida. Sin algún sistema de calificación es imposible saber si las conductas, en forma de actitudes, valores, habilidades, conocimientos, etc., han sido inculcadas, inhibidas o alteradas.

Evaluación, valoración y medición

En este sector, los tres términos, *evaluación*, *valoración* y *medición*, se utilizan a menudo indiscriminadamente y como si fuesen intercambiables. Para mayor claridad vamos a intentar distinguirlos. La evaluación, el término más amplio, incluye no sólo el proceso que determina cuáles son los

resultados educativos reales comparándolos con los esperados, sino que implica además el juzgar si los cambios efectuados son deseables. Los juicios de este tipo, evidentemente, se basan directamente en los objetivos educativos, pero no podrán emitirse hasta que no se lleve a cabo la valoración. Aunque se puede juzgar *a priori* si un resultado esperado es deseable o no —es de suponer que se tendrán en cuenta todos los datos relevantes— los juicios sobre la eficacia y las experiencias elegidas, el contenido utilizado y la manera de utilizarlo dependen enteramente de algún tipo de valoración del cambio de conducta. La valoración es un prerrequisito de la evaluación. Las preguntas: "¿Ha habido cambio?, y si es así, ¿cuánto, y en qué sentido?" son cuestiones a las que la valoración ha de responder. "¿Era adecuada la experiencia educativa para este propósito?" es una cuestión de evaluación, puesto que implica un juicio. Si se han utilizado más de un conjunto de experiencias y de material con varios grupos equivalentes, la valoración podrá responder a las preguntas: "¿Existe alguna diferencia entre los grupos?" Si es así, "¿cuánta y en qué sentido?". Pero a la evaluación se le pide que conteste a la problemática de: "¿Qué conjunto (o método) es el mejor?"

La evaluación, por lo tanto, implica emitir un juicio con respecto a ciertos criterios. "Valoración" será el término que designe el proceso de investigar el nivel de un determinado grupo, normalmente en relación con las conductas esperadas. Se trata, por lo tanto, de un término general que incluye el más restringido de la medición, el cual implica un cierto tipo de escala en la que los individuos se ordenan de acuerdo con lo que saben o pueden hacer.

En cuanto a las calificaciones, hay que recordar que las escalas educativas son del tipo conocido como "ordinal". Aunque el orden de posición en la escala es el mismo que el orden de magnitud numérica, los intervalos de la escala no son iguales. Aunque es relativamente fácil medir los cambios en la estatura de un alumno, en su peso, en su capacidad pulmonar, en su velocidad al correr o en su fuerza al apretar, tanto el proceso como la escala de medida son completamente diferentes para los demás conductas que con-

WHEELER, "La evaluación" en: El desarrollo del currículum escolar. España, Santillana 1985, pp. 287-308.

las propiedades básicas de las escalas educativas y los su-
puestos que son relevantes para la medida educativa, y de-
ben emplear técnicas estadísticas apropiadas cuando abor-
dan el tema de la valoración.

El fin de la fase de evaluación es proporcionar datos sobre
el carácter, el sentido y la medida de los cambios de con-
ducta provocados por los esfuerzos educativos, y utilizar esta
evidencia como guía para modificar cualquier fase del pro-
ceso del currículum. La única evidencia aceptable a este
respecto es la del cambio que se opera en el alumno. En
algunos casos será posible medirlo cuantitativa y objetiva-
mente; en otros puede bastar con algún tipo de valoración cua-
liativa. Además de valorar los cambios de conducta de los
individuos o de los grupos, es necesario juzgar los objetivos
de la escuela y la idoneidad de las experiencias de apren-
dizaje, el contenido y la organización, y los métodos de
enseñanza utilizados para alcanzar dichos objetivos. Esto
significa que, al menos, hay que tener en cuenta lo siguiente:
la valoración de la conducta relevante del alumno, y los fac-
tores ambientales que puedan influir en dicha conducta,
el efecto del programa de evaluación en el aprendizaje y
en las motivaciones del alumno, y la evaluación de todas
las fases del proceso del currículum. Evidentemente los
tres primeros puntos se basan directamente en el cuarto.
Ya debe estar claro que el proceso del currículum detallado
en este volumen requiere una evaluación cada vez más am-
plia y compleja. La conducta es más difícil de medir que el
conocimiento, pero la conducta es el resultado esperado de
la educación. Para empezar, la evaluación que se requiere
ha de ser continua, coherente y comprensiva.

La continuidad

Tanto si el programa de evaluación está encaminado a me-
jorar el curso, como si lo está a tomar decisiones sobre los
individuos, o a regular la administración (1), en cualquier caso,
ha de ser continuo y ha de constituir una parte integral de la
enseñanza en clase. El paradigma en este caso es de valo-

ración continua. Si al adelantarse un paso
hacia una meta se hace una valoración, ésta puede servir
para modificar o volver a valorar la meta, así como para to-
mar decisiones sobre el siguiente paso a dar. Este ciclo de
pasos hacia delante y hacia atrás puede repetirse hasta al-
canzar la meta total, tanto si se trata de la que se estableció
originalmente como si ha sido modificada a la luz del *feed-
back*. Sin ello, el alumno carece de información sobre sus
propios progresos hasta finalizar la secuencia, y puede sa-
car conclusiones falsas, habituarse a actitudes equivocadas
o aprender falsos métodos de trabajo. El *feedback* es, asi-
mismo, importante para la motivación, puesto que la meta
está periódicamente sujeta a nuevas valoraciones, en tér-
minos de ser más o menos deseable y factible. Puesto que
los datos objetivos muestran que la meta es razonable,
habrá que comprometerse con ella, al máximo. En esencia,
este procedimiento permite al alumno tomar parte en la
definición de metas más realistas. Esto es importantísimo,
puesto que las únicas metas de aprendizaje eficaces son
las que el alumno se propone por sí mismo, o las que el
profesor fija y el alumno está dispuesto a aceptar como
propias.

En lo que respecta al desarrollo de los cursos, o unidades
de trabajo, es válido el mismo principio. Hay que buscar
pruebas de que la unidad va progresando adecuadamente
durante el desarrollo de la misma, y no cuando se termina,
porque entonces es lógico que los programadores se mues-
tren reacios a desechar un material conseguido a costa de
tiempo y trabajo. En cada paso que se avanza hacia la cons-
trucción de una unidad existe algún tipo de valoración, y
debe someterse a una evaluación continua.

La coherencia

Es posible que los objetivos del currículum presenten un
cierto tipo de jerarquía, en cuanto que pueden disponerse
por orden de importancia. Esa misma jerarquía debe estar
presente en los diversos aspectos de la evaluación. Si se
consideran más importantes los conceptos que los hechos,

habrá que poner mayor énfasis en la evaluación del desarrollo de los conceptos que en el recuerdo de hechos específicos. Si se da más importancia a la medición de los resultados cognoscitivos es porque se piensa que son más importantes, y no porque su medición sea más fácil u ofrezca más confianza. Análogamente, si el principio de organización utilizado es el de los grandes campos, o el de los núcleos, entonces los tests no deben ser los mismos que se han desarrollado para un programa por materias separadas.

La comprensividad

Nunca se subrayará bastante la necesidad de un programa de evaluación comprensivo. Existe un paralelismo necesario entre objetivos explícitos y su evaluación. La valoración es tan necesaria para la conducta inicial como para los resultados finales, puesto que sin ella es imposible saber si se ha operado algún cambio en la conducta. Por consiguiente, hay que aceptar la idea de que todos los objetivos escolares han de ser valorados de alguna manera. A pesar de lo inadecuado de los dispositivos de mediación para algunas categorías de la conducta, debe tratarse de valorar un espectro de conducta de igual amplitud al que indican las metas y objetivos de la fase 1. Dicho con más detalle, debe hacerse:

1. Una valoración del status de los niños en las variables importantes para: a) conocer su situación actual; b) poder expresar los objetivos generales en términos de conducta específica, adecuados para esos niños; c) escoger experiencias, contenido y métodos adecuados y lograr una integración tanto horizontal como vertical.
2. Durante cualquier secuencia educativa determinada: a) la evaluación de los progresos hacia la meta, y b) la modificación de las metas, experiencias, contenido y métodos, de acuerdo con la teoría del *feedback*.
3. Al final de cualquier secuencia determinada, la valoración de los resultados reales para ver: a) lo que se ha aprendido, y b) en qué medida se han integrado esos aprendizajes con los de otras experiencias de aprendizaje.

La naturaleza de alguna de estas importantes variables se indica en la siguiente lista:

- Modalidades que se utilizan para el aprendizaje. Estructura de la motivación.
- Nivel de aspiración. Experiencias ambientales.
- Intereses. Nivel de madurez.
- Características de la personalidad importantes para el aprendizaje.
- Ambiente sociocultural.
- Relaciones con los demás (incluido el grupo).
- Adaptación social, habilidades y actitudes.
- Rigidez, prejuicio y actitud hacia la autoridad.
- Creatividad. Valores y ética.
- Salud física, conocimiento, actitudes y valores.
- Conocimiento estético, habilidades y capacidades.
- Conocimiento cognoscitivo, habilidades y capacidades.
- Habilidad para comunicarse. Habilidad para el estudio.
- Habilidad para el trabajo.

Como el niño en la escuela forma parte de un grupo (o una serie de grupos), es preciso también saber algo acerca de las metas del grupo, el movimiento hacia las metas, el liderazgo, la estructura y el proceso del grupo con respecto a la clase y a otros grupos de la escuela.

Los principales medios de valoración y medición serán: a) tests y baterías de tests uniformados (Inteligencia, logros, disposición, diagnóstico, pronóstico); b) tests y exámenes hechos por el profesor; c) observación y registro sistemático; d) escalas de medición; inventarios, cuestionarios y cártalos; e) preguntas abiertas, completar frases y técnicas para hacer proyectos simples; f) sociometría.

La objetividad

Puede afirmarse con seguridad que la medición es más objetiva, válida y digna de confianza en el sector cognoscitivo que en el sector afectivo, y más a nivel cognoscitivo bajo que a nivel alto. Esto significa que hasta que no se desarrollen tests que midan de manera satisfactoria los aprendizajes.

más complejos en el dominio afectivo y psicomotor, al igual que en el cognoscitivo, habrá que seguir dependiendo de la valoración subjetiva de los especialistas competentes, así como de los instrumentos objetivos. Pero hay que recordar que subjetividad/objetividad es un *continuum*, no una dicotomía. Los datos sobre la valoración en educación varían desde los que son principal, aunque no enteramente, subjetivos hasta los que son de igual modo objetivos. El fin ha de ser el conseguir que cada instrumento de valoración sea lo más objetivo y digno de confianza posible, de manera que pueda utilizarse como prueba fidedigna, en el proceso de evaluación. Al implicar juicios de valor el proceso ha de ser necesariamente más subjetivo. Lo objetivo, lo que hay que evitar es adscribir a un dato más objetividad, confianza o validez de la que realmente tiene, lo que es fácil que ocurra con las mediciones expresadas en escalas numéricas, tales como los resultados de exámenes y las puntuaciones de los tests de realización unitarios. Naturalmente, la evaluación exige toda la evidencia disponible. Si ésta es cuantitativa y objetiva, mucho mejor; pero es preciso advertir que tanto los datos como los supuestos en los que se basan deben ser analizados cuidadosamente.

La validez

La validez ha sido mencionada como criterio importante en las fases 2, 3 y 4. Análogamente, se encuentra en el centro de la relación entre la fase 1 y la 5. El asunto requiere ser tratado con más detalle de lo que es posible aquí (2); pero lo que sí se puede decir es que hay que tener buen cuidado de estar seguro de que todas las mediciones empleadas tienen validez empírica, en el sentido de relevancia (estrecha coincidencia entre lo que mide el test y la función que se supone ha de medir).

Los tests uniformados que más se utilizan dentro de las materias señaladas en los niveles bajos del Conocimiento (clase 1.00), en general alcanzan bastante altura tanto en lo que se refiere a validez como a fiabilidad (la precisión y la

coherencia con las que un test mide lo que mide en el grupo que mide). En cuanto a otros objetivos, debe prestarse mucha atención a las valoraciones, pues muchos de los objetivos de nuestras categorías 1c, 2 y 3a no se cuidan de forma adecuada. Evidentemente, los profesores tendrán que elaborar sus propios instrumentos, y en esta labor les puede servir de gran ayuda la *Taxonomía*.

No hay que olvidar, sin embargo, que no basta con medir el conocimiento de los hechos y las habilidades fundamentales. La capacidad intelectual superior no se mide con simples tests sobre los hechos, puesto que éstos se refieren solamente a los niveles inferiores de la clase 1.00 (Conocimiento). Lo cierto es que el tipo de instrucción que lleva al Conocimiento basado en el simple recuerdo puede incluso interferir en la consecución de ciertos resultados educativos importantes. De la misma manera, los tests elaborados para medir los resultados cognoscitivos no demuestran necesariamente que se hayan logrado los resultados afectivos apropiados. Según esto, por tanto, los procedimientos usados deben medir realmente las conductas definidas operativamente en la fase 1.

Dado que actualmente a lo que se da más importancia es a la evaluación de lo que se ha conseguido, es muy probable que en lo que se haga mayor hincapié sea en la validez del contenido, es decir, en el grado en que los tests dan una idea adecuada de los tipos de situaciones relacionadas con los objetivos educativos.

En la práctica, dada la dificultad de medir las características de las situaciones, la validez del contenido se limita a indicar hasta qué punto un test es una prueba representativa de la materia tratada. Si las anteriores etapas del proceso del curriculum no se han realizado adecuadamente, o si la gradación de la importancia de los objetivos específicos (especialmente en la tabla de especificación-tests), no es la misma que la utilizada en el proceso de instrucción, entonces, aunque el test sea válido respecto a la clase, puede no serlo en lo que respecta a los fines generales de la escuela. Ya hemos señalado esta posibilidad, pero quizás valga la pena subrayar que la validez del contenido está inseparable-

biamente ligada a los objetivos educativos, a su gradación de importancia y al carácter de la tabla de especificaciones derivada de ellos.

La validez concurrente, o de nivel, se refiere a la proporción en que la conducta elicited por el test corresponde a algún criterio externo que existe en aquel momento en el repertorio de conductas del alumno. Excepto en cuanto que trata de medir la conducta en su forma ordinaria, difiere poco de la validez predictiva, la cual está relacionada con el grado en que la puntuación del test está asociada con la conducta futura. El hecho de que un test pueda medir satisfactoriamente la conducta real, aunque no valga como predicción, es prueba de la importancia de los factores temporales en la evaluación. La relación ya establecida entre las metas últimas y las próximas indica que los objetivos específicos y estar directamente relacionados con las metas últimas. Siempre que se encuentre un test que tiene validez respecto a un nivel superior y poca validez predictiva dentro del mismo campo, es evidente que no se han mantenido las condiciones necesarias para el cambio de conducta, o que se ha llegado a conclusiones erróneas sobre las relaciones entre metas últimas y próximas o entre metas y experiencias, contenido, organización o metodología. Esto indicaría que hay que tener en cuenta la evaluación durante un período más largo de lo que generalmente se considera necesario.

El hecho de que diferentes cursos puedan producir muchos de los mismos resultados generales constituye una razón más para valorar el desarrollo de la educación general de los alumnos a lo largo de un período de tiempo más largo. Los objetivos complejos en los grados más altos, tanto del campo cognoscitivo como del afectivo, probablemente no se alcanzarán más que a través del período completo de escolaridad. La valoración de estos objetivos exige un plan de evaluación que cubra más de un grado, e implique a más de un profesor y a más de un curso, pues el avance hacia objetivos complejos sólo puede medirse mediante una valoración longitudinal a lo largo de un período de tiempo bastante considerable.

Dado que no existen cursos determinados que se refieran directamente a los niveles superiores de las clasificaciones taxonómicas que se supone que los alumnos han de adquirir a lo largo de una serie de cursos, estos objetivos complejos y a largo plazo nunca aparecen en los exámenes de las asignaturas. Además, los objetivos específicos de la instrucción en las materias individuales no responden colectivamente de las metas últimas del currículum total. Por consiguiente, es muy necesaria la descripción comprensiva del desarrollo educativo total de los individuos y de los grupos. Las medidas del progreso individual exigen mediciones repetidas de los mismos rasgos, o la administración reiterada de los mismos tests, o de tests comparables, lo cual, generalmente, no se lleva a cabo.

En cuanto a la validez, el tipo más importante es la validez del constructo, que se demuestra poniendo de manifiesto que las puntuaciones de los tests pueden servir para hacer inferencias coherentes con la teoría en la que se basa el test. Puesto que ésta teoría se deriva de los mismos cuerpos de doctrina que hemos utilizado para programar el proceso del currículum, habrá que prestar mucha atención a la validez del constructo.

Es preciso subrayar dos puntos. Primero, la dificultad de encontrar mediciones independientes y satisfactorias de los atributos que se suponen valorados por los instrumentos de medición da generalmente como resultado que no tengan más que una validez aparente. Y aún más importante, no es inconcebible que los tests que tienen una gran validez concurrente o de contenido respecto a las metas próximas o a los objetivos específicos tengan poco valor para pronosticar las metas últimas. La evaluación no sólo se refiere a los cambios de conducta absolutos y diferenciales en una unidad de trabajo, sino a las relaciones entre estas unidades (relación tanto vertical como horizontal) y al grado en que la organización de las experiencias con contenido produce efectos acumulativos asociados con las metas últimas de la educación. Los análisis de las tendencias y los estudios de las secuencias proporcionan datos sobre los cambios de conducta de los individuos y de los grupos, y sobre lo adecuado.

o no, de las operaciones seguidas en las fases 2, 3 y 4. Evidentemente pueden desempeñar un importante papel en la fase 5.

Participación en la valoración

Históricamente, los que aprendían y los que ejercían la enseñanza han constituido dos clases separadas: los alumnos y los profesores. Al aumentar los medios educativos se ha producido un incremento en el número de personas relacionadas con la evaluación, especialmente con la valoración y la medición. Existen inspectores o superintendentes que pueden valorar tanto a los profesores como a los alumnos, directores o rectores (que pueden realizar la misma función que los inspectores) y tribunales de exámenes: externos (que dirigen los exámenes de bachillerato, del School Certificate, del Regents' y otros semejantes). Como se observará, todas estas personas se encuentran en el polo docente del proceso educativo, pero aparentemente no existe en ninguna parte un lugar para el otro polo; el alumnado. Sin embargo, en la vida cotidiana, los niños de las escuelas están constantemente midiendo, evaluando y valorando los cambios de conducta que observan en sí mismos y en los demás. Es cierto que a través del proceso de concesiones mutuas de los métodos tradicionales del aula se permite a veces que sean los niños los que califiquen su propio trabajo o el de los otros. Pero, incluso entonces, el profesorado com- prueba el trabajo. Esta situación se produce a causa de los roles de profesor y alumno. En un aspecto de su rol, al profesor se le considera como un juez; es asunto suyo señalar hasta qué punto ha habido un cambio de conducta, especialmente en el campo cognoscitivo. No se reconoce, sin embargo, que en el rol de alumno pueda ser éste quien valore su propio cambio educativo o el de sus compañeros, aunque ambos tipos de valoración son frecuentes fuera de la escuela.

Puesto que la evaluación es parte integral del proceso de aprendizaje, es preciso contar con el alumno, en la medida en que sea posible y él sea capaz de hacerlo, a la hora de evaluar su propio aprendizaje. La única manera de saber si

ha cambiado la conducta es comparando las valoraciones de nivel, tanto antes como después de la experiencia de aprendizaje. El aprendizaje es un proceso activo, y una parte de esa actividad es comprobar los resultados. Con demasiada frecuencia esta comprobación de los resultados de las experiencias educativas queda aislada de las propias experiencias. Y es sobre todo muy probable que ocurra así si se la considera como una prerrogativa del profesor. Entonces, si él no la realiza, no se lleva a cabo, con lo cual la experiencia de aprendizaje queda incompleta. Especialmente durante los últimos años de escuela, los niños necesitan tomar conciencia del proceso de aprendizaje y es preciso que se les anime a aprender y a habituarse a valerse de métodos eficaces para la resolución de sus problemas. La mayoría de las escuelas aceptan implícitamente las posibilidades de transferencia, sin preocuparse mucho por la evidencia experimental, pues los profesores esperan que sus alumnos aprendan en la escuela una serie de procesos que transferirán a las situaciones reales. Pero que esto suceda es poco probable, a no ser que se les señalen explícitamente las generalizaciones, conexiones y aplicaciones. Si los alumnos no tienen conciencia de cómo se aprende y bajo qué condiciones se alcanza al máximo, la transferencia será escasa. El estudiante que se ha habituado a que alguien le examine no ha aprendido a organizar su propio aprendizaje.

Los tests proporcionan los elementos de juicio para decidir si se avanza o no hacia una meta. Para el individuo, el someterse continuamente a los tests (esto es, apelar a la realidad) resulta esencial. El aprendizaje se lleva a cabo con más éxito en unas condiciones que permitan que se retengan las respuestas correctas y en las que el feedback cognoscitivo o el inmediato conocimiento de los resultados pueda establecer la probabilidad de un refuerzo efectivo. La evaluación continua, por parte del alumno, sirve para que el establecimiento y la consecución de las metas sean más efectivos, con la probabilidad de que la motivación se haga continua o incluso que vaya en aumento.

Otro punto que hay que considerar es la relación entre las metas de progreso direccional y las metas inmediatas o a

corto plazo. El marco de referencia del profesor estará relacionado probablemente con las metas a largo plazo, y el de los alumnos, con las metas a corto plazo. El proceso de valoración debe ayudar a aclarar los objetivos, de manera que la participación de los alumnos en la valoración llegará los propósitos a corto plazo de los alumnos con las metas a largo plazo del profesor. El comprobar continuamente la realidad no sólo mejora el aprendizaje individual; también mejora la estructura de grupo, lo que a su vez ayuda al individuo. El punto de vista de que la valoración es parte continua e integral del desarrollo del currículum y exige la participación de los alumnos, no siempre concuerda con las ideas tradicionales sobre los tests. Desde el punto de vista tradicional, los tests y los exámenes se hacen para motivar a los alumnos en sus tareas y para mantener ciertos niveles medios (cualquiera que sea el significado de esto). Incluso hoy día muchos profesores sostienen, implícita o explícitamente, opiniones tales como "si no hay examen, los chicos no estudian" (lo cual no es cierto), y "si no hay exámenes, no se pueden mantener los niveles medios" (lo cual no tiene sentido).

Se puede considerar, sin embargo, la posibilidad de que los estudiantes no sólo sean examinados por el profesor o alguna otra persona adulta de la jerarquía docente, sino que tomen parte en la planificación, construcción y organización del dispositivo de valoración y en la evaluación de sus progresos diarios. Uno de los fines de la educación general puede ser la internalización de las normas de la comunidad; el punto de vista tradicional subraya solamente la necesidad de la obligatoriedad externa. Efectivamente, las complejas sociedades modernas precisan sanciones formales y organismos que las administren, pero, de hecho, estos organismos son menos eficaces para promover una conducta correcta que los valores internalizados y las costumbres. Resulta triste el comentario de un reciente informe sobre la internalización de los valores educativos que afirma que los exámenes (en el contexto se trata de exámenes externos) "han llegado a alcanzar tal importancia para los ingleses que las materias que no se someten a examen se miran con indiferencia." (3).

En lo que respecta al profesor individual, su participación en la selección y elaboración de los dispositivos de valoración y en el programa de evaluación probablemente mejore la enseñanza, puesto que a través de ella se hacen más claros y significativos las metas y objetivos. Cuando un profesor tiene que elaborar algún "test", esto debería forzarle a revisar de nuevo sus objetivos. La relación adecuada entre las fases 1 y 5 —toda meta indicaba en la fase 1 debe ser tasada de alguna manera— presenta una doble ventaja: evita que las formulaciones de metas y objetivos se llenen de materia inservible e inútil, y hace recaer en el profesor la responsabilidad de valorar los objetivos formulados.

En los casos en que el currículum pueda ser centralizado y el material utilizado sea en su mayor parte idéntico para todos los profesores de un sistema, será conveniente que los tests sean elaborados por ese mismo personal central y que les sean devueltos para calificarlos. Dado que los resultados son multidimensionales, habrá que intentar comprender de alguna manera, a este nivel, cuáles son los factores que conducen al éxito o al fracaso y bajo qué condiciones operan dichos factores. Este tipo de análisis puede servir de ayuda para entender el proceso de aprendizaje y para aclarar los determinantes de la conducta, fortaleciéndose de este modo el cuerpo de doctrina del que depende el desarrollo del currículum.

Valoración y aprendizaje

La medición y otros tipos de valoración juegan evidentemente un papel importante en la motivación del aprendizaje y en el propio aprendizaje. La función activadora de los exámenes se pone de manifiesto en el fenómeno de "empollar" que se da en algunas escuelas y sistemas escolares. La función directiva también está clara. Los estudiantes tienden a dar importancia a aquello de lo que van a examinarse, mientras que los profesores acentúan los objetivos que van a ser comprobados por otros o que ellos mismos pueden comprobar más fácilmente. "Cuando las escuelas, los profesores y los alumnos se valoran hasta cierto punto por los

resultados de los tests, la índole del test determina en gran medida la calidad del proceso escolar, lo que los profesoras enseñan y cómo lo enseñan—lo que los alumnos aprenden y cómo lo aprenden—, (4). Además, el tipo de examen influye muchísimo en la forma en que los alumnos estudian para prepararlo.

El peligro de esto radica en que el tipo de procedimiento utilizado para la valoración pueda influir en las otras fases del proceso del currículum, y en que se preste mayor atención a los objetivos que puedan comprobarse con mayor facilidad. Esto sería invertir la relación entre las fases 1 y 5. En todos los casos son las metas y objetivos los que deben determinar la valoración, y no la facilidad de la valoración la que determine los objetivos. Desgraciadamente, esto es lo que ocurre a veces en relación con los exámenes, sobre todo, los exámenes por libre.

Los exámenes

Aunque todos los expertos en mediciones están de acuerdo en que hay que aclarar cuáles son los objetivos y que éstos deben exponerse claramente antes de hacer la valoración, uno de los inconvenientes de casi todos los programas o cursos —y en esto es en lo que se basan la mayoría de los exámenes— es que en ellos no se formulan los objetivos, sino que simplemente se detalla el contenido. Si los objetivos no están formulados claramente, no se puede evaluar, pues es imposible hacer una valoración sin saber lo que hay que valorar. Para remediarlo, el que valora se ve forzado a establecer unos objetivos. Sin embargo, en esta fase no es tarea del educador proporcionar objetivos. Su labor es valorar y emitir juicios sobre las conductas observables y sobre la organización e integración de experiencias y contenido que han desembocado en dichas conductas. Si en la fase de valoración se ve obligado a hacer algo que debería haberse hecho en una fase anterior, no puede esperar más que encontrarse con una serie de dificultades.

Puede servirnos de ejemplo el caso de una sencilla valoración de la conducta cognoscitiva (intelectual) en un sector

aislado, que es una tarea que se proponen todos los tribunales de exámenes. Para valorar el Conocimiento (clase 1.00) solamente, es preciso especificar por lo menos 18 grupos de objetivos, dos por cada uno de los apartados siguientes: (1.12), terminología (1.11), convenciones (1.21), tendencias y secuencias (1.22), clasificaciones y categorías (1.23), criterios (1.24), metodología (1.25), principios y generalizaciones (1.31) y teorías y estructuras (1.32). Para cada grupo hay que especificar la selección hecha entre todo el Conocimiento posible del sector y el grado de precisión con que debe conocerse. Si se va a valorar la Comprensión (clase 2.00), las tres clases, de traducción, interpretación y extrapolación, pueden cada una enfocarse en relación con las nueve subclases del Conocimiento citadas anteriores. La Aplicación (clase 3.00), el Análisis (clase 4.00) o la Síntesis (clase 5.00) pueden aplicarse por separado al Conocimiento previamente comprendido, mientras que la Evaluación (clase 6.00) requiere que se combinen procesos necesarios de cualquiera, o de varias, de las cinco categorías anteriores. Una valoración adecuada sólo es posible si se detallan los objetivos específicos en relación con todas las categorías de una materia determinada. A no ser que el profesor y el examinador hayan realizado análisis semejantes, es probable que el examinador trate de valorar teniendo en cuenta objetivos que el profesor no ha considerado, o a un nivel o con un grado de precisión diferentes a los del profesor. Puede incluso ocurrir que el examinador no intente valorar la consecución de los objetivos en los que el profesor ha gastado gran parte del tiempo. La mayoría de las discusiones entre profesores y examinadores se debben a la falta de formulaciones de objetivos. El profesor puede haberse centrado en el Conocimiento (clase 1.00), mientras que el examinador exige Comprensión (clase 2.00), o que el objetivo del profesor sea la Comprensión, mientras que el examinador exige Aplicación (clase 3.00) o Evaluación (clase 6.00). Si el profesor y los alumnos opinan que el examen es demasiado difícil, esto se debe con frecuencia a que existe un desacuerdo en cuanto al nivel y al grado de precisión del Conocimiento que se exigen.

Los peligros anteriormente indicados no son inherentes a los tests y exámenes, que sólo son instrumentos, sino que

dependen del uso que se les dé. Teniendo en cuenta que algunos de los más importantes objetivos de la educación son extremadamente difíciles de medir en una escala cuantitativa, los objetivos simples que se pueden medir fácilmente (número de hechos recordados, etc.) pueden aceptarse como sustitutos de los objetivos más difíciles de valorar. A este proceso le ayuda la seudoprecisión de la calificación media. Evidentemente, los conceptos de validez son vitales para toda cuestión relacionada con los exámenes.

Además, a veces se da excesiva importancia a los exámenes porque no se distingue entre evaluación y medición. Un examen es un instrumento de medida, un dato del que hay que valerse en la evaluación. Si es el único dato que se tiene en cuenta, existe la tendencia a utilizarlo para unos fines para los que no es necesariamente adecuado. Esto no quiere decir que los exámenes hayan de ser abolidos. Los exámenes juegan un importante papel en la evaluación, pero no deben usarse para fines para los que no están concebidos, ni pensar que van a dar una información que no proporcionan. Son instrumentos de valoración como otros cualesquiera; si se les utiliza adecuadamente y se reconocen sus limitaciones, proporcionan una información útil sobre la escuela. Pero no son los únicos instrumentos de valoración y, sobre todo, no son los que fijan los objetivos educativos.

Al tratar de los exámenes surge la cuestión de los niveles medios. Este término se utiliza corrientemente con dos sentidos diferentes por lo menos. Los niveles medios de los examinadores son, por lo general, lo que ellos estiman que los candidatos deben hacer; y mantener el nivel significa hacer el examen difícil o no dejar pasar demasiados candidatos. Si los objetivos no están definidos muy claramente y los exámenes no comprueban todos esos objetivos de manera válida y digna de confianza, puede haber diferencias de opinión sobre esos niveles medios. La amplia evidencia experimental muestra claramente que puede haber diferencias de opinión sobre los niveles medios de los examinadores que, en el caso de un mismo examen, no estén de acuerdo sobre los niveles medios de los exámenes, y que, además, el mismo examinador dé distintas puntuaciones a un mismo examen en diferentes ocasiones.

Los niveles medios son los niveles deseables (o deseados) de consecución y que deben expresarse operativamente como resultados de la enseñanza. Constituyen, por lo tanto, metas u objetivos. Aunque las medidas válidas y fidedignas puedan indicar si las conductas esperadas han sido conseguidas realmente, los juicios sobre los niveles medios son más importantes para cualquier fase del proceso que para la fase 5, excepto en el caso en que los niveles medios se hayan evidenciado, en los procedimientos de valoración, como objetivos específicos inadecuados. La fase de evaluación se refiere a las conductas reales, no a las esperadas (o niveles medios).

La evaluación y las otras fases del currículum

Es preciso, finalmente, considerar la evaluación como una fase formal en el proceso del currículum. La última fase completa el ciclo, puesto que requiere un nuevo estudio de las metas y objetivos a la luz de la información que suministra. Esto permite que se repita el ciclo del currículum en forma diferente. La selección de objetivos y la valoración de hasta qué punto han sido alcanzados son fases complementarias, ninguna de las cuales puede originar un currículum satisfactorio sin la ayuda de la otra. Es imposible elaborar un test válido sin conocer clara y detalladamente los objetivos, de manera que para conocer la naturaleza y la amplitud del cambio de conducta es esencial saber exactamente cuál es el comportamiento que hay que medir. Todo esto significa que los objetivos del currículum deben definirse adecuadamente, ser posible en términos de conducta. De ello se desprende, pues, que incluso cuando se detallan objetivos por vez primera sería preciso tener algo en cuenta la valoración. Si se desea un cambio, ¿cómo se medirá? Si no hay manera de valorar el progreso hacia un determinado objetivo, quizás fuese mejor reemplazarlo por otro que sea igualmente deseable, pero que pueda ser valorado.

Una vez que se han dejado atrás las generalidades, la valoración desempeña un papel muy importante en la selección de metas próximas y de objetivos específicos. Ayudar a los ni-

nos a pensar con lógica y claridad, o enseñar a los niños cómo se aprende, o formar hombres y mujeres tolerantes con las diferencias, no requiere, para que se seleccionen como fines convenientes contar necesariamente con datos de valoración. Sin embargo, tan pronto como se intenta convertir cualquiera de ellos en una formulación específica y operativa de importancia para un determinado grupo de niños de una cierta edad en una clase, se hacen necesarios los datos. Un tipo de datos se refiere a las características de esos individuos en particular. Tales datos solamente se consiguen mediante la valoración. De manera que, en la práctica, la valoración de las variables obvias es necesaria incluso en la fase 1.

Aunque el proceso de la valoración de la fase 5 pueda, mediante los procedimientos de evaluación que le siguen, dar lugar a una elaboración o modificación de los objetivos detallados en la fase 1, existe otra forma más general de relación entre la valoración y la selección de los objetivos del curriculum. La respuesta a la pregunta «¿qué es lo más valioso de nuestra sociedad, rápidamente cambiante, para los niños y para la juventud?», se da demasiado a menudo en términos de una orientación «filosófica» y de la práctica tradicional. Cuanto más se sabe del desarrollo humano y de la interacción social, más deben deducirse los objetivos de los hechos, principios y teorías relacionados con estos sectores. La conducta, el proceso de aprendizaje, las necesidades sociales e individuales, todo ello da lugar a modelos teóricos. Lo fundamental en la elaboración de un modelo es la observación científica, controlada, cuyo principal componente es la medición, o al menos la valoración, del cambio. No conseguiremos ningún curriculum adecuado más que si nos vamos de la valoración, pues ésta es la manera de conseguir mejores modelos de la sociedad o la conducta humana.

Una parte importante del proceso enseñanza-aprendizaje es la provisión de experiencias y contenido convenientes para alcanzar los objetivos. La conveniencia depende de una serie de factores, entre los cuales están la experiencia previa y la organización de la conducta del alumno en particular: sus intereses, actitudes, valores, capacidad y conocimiento. Para

poder adaptar las experiencias de aprendizaje a la capacidad, intereses y aprendizaje previo, el profesor debe saber cuáles son éstos; además, debe tratar de valorar los propósitos del alumno, su estructura motivacional y su grado de disposición. Esto es igualmente válido para las metas mediatas y próximas y para los objetivos específicos. Como la clase de un año es diferente de la del año siguiente y los niños son todos ellos diferentes entre sí, es preciso conocer todas estas diferencias para seleccionar las experiencias y el contenido.

Así, la función de la valoración en las fases 2 y 3 es, en cierto sentido, doble: proporciona datos que ayudan a seleccionar las experiencias y el contenido, e indica la disposición por parte del individuo o del grupo. Havihurst ha utilizado la expresión «momento propicio para enseñar» para referirse a aquel en el que el individuo está dispuesto y la sociedad lo requiere. La valoración ayuda a identificar tales momentos propicios. En clase, el profesor, como representante de la sociedad, puede modificar sus exigencias hasta que la valoración indique que los individuos están realmente preparados para el aprendizaje requerido. En muchos sectores hay ciertas experiencias y ciertos contenidos que son requisitos previos para seguir avanzando, así que la valoración se convierte en determinar tanto el tipo de requisitos previos como el momento en que los individuos han alcanzado ciertas etapas o niveles.

Análogamente, la organización e integración de las experiencias y el contenido pueden ser mejoradas mediante la valoración y la evaluación. Como en las fases 2 y 3, la fase de organización implica una cierta elección, pues hay pocas experiencias educativas de un carácter tal como para que sólo se pueda dar un conjunto de principios de organización. La valoración ayuda a determinar los efectos de un tipo de organización en comparación con los de otra, así como la manera en que la organización a un nivel pueda influir en la integración con otro.

A lo largo de un cierto periodo de tiempo, el proceso de valoración puede ser comparativo. Dada la posibilidad de que se hagan diferentes planteamientos de los mismos objetivos,

el profesor puede intentar alcanzar los mismos fines de maneras muy diferentes con clases distintas, tanto si son del mismo como de distinto nivel. Se podrán entonces comparar los diferentes efectos de dos o más planteamientos. La comparación implica una valoración de las diferencias y una valoración de los métodos en relación con los criterios.

No se puede negar que este tipo de enfoque experimental de la educación es deseable, ni que es necesario que el desarrollo del currículum esté basado en principios coherentes. La investigación específica dentro del aula es una manera de alcanzar estas aspiraciones. No es posible obtener evidencias sobre la validez de las hipótesis que cimentan los procedimientos de clase, o sobre la eficacia de la integración de los distintos niveles, más que a través de una adecuada evaluación.

La relación entre la valoración y la fase 4 es muy parecida a la que hay entre la valoración y la fase 1. Para elaborar algún test que demuestre si se han alcanzado los objetivos es preciso que se detallan estos objetivos. De la misma manera, para evaluar los elementos de organización de las experiencias educativas es necesario, en primer lugar, establecer cuáles son estos elementos y en qué principios están basados. La información sobre la eficacia de la organización es directamente relevante para las hipótesis sobre los principios en los que se basa y puede utilizarse para comprobar si dichas hipótesis son coherentes. El procedimiento es esencialmente el mismo que en la fase 1. Al valorar la naturaleza y amplitud del cambio de conducta es posible evaluar las hipótesis iniciales y, o bien modificarlas, o bien seleccionar nuevos principios de organización que sean más apropiados para comprobarlas.

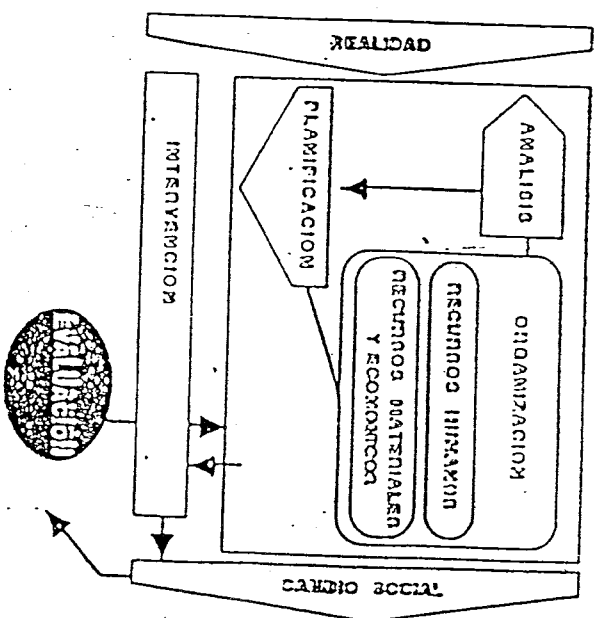
Los detalles y métodos necesarios para una evaluación satisfactoria pueden encontrarse en los trabajos técnicos sobre el tema. Hay que señalar, sin embargo, que la fase de evaluación es una de las más importantes de todo el proceso del currículum, y que la evaluación y la valoración son partes integrantes de todas las fases. Afirmaciones tales como «el proceso de valoración de un programa escolar determinado es una parte del currículum, de la misma forma que pueda

serlo el aprendizaje del alumno» (5) olvidan, entre otras cosas, la naturaleza del proceso del currículum. El punto de vista que nosotros mantenemos es el mismo que expresa Cronbach: «La evaluación es una parte fundamental del desarrollo del currículum, no un apéndice» (5).

Notas y referencias.

- (1) Ver Cronbach, L. J., «Evaluation for course improvement», en Heath, R. W. (ed.), *New Curricula*, Nueva York: Harper and Row, 1964, p. 232.
- (2) Para un análisis detenido de la validez, véase Cureton, E. E., «Validity», cap. 16, en Lindquist, E. F. (ed.), *Educational Measurement*, Washington, D.C.: American Council on Education, 1951, pp. 621-694; Cronbach, L. J., y Meehl, P. E., «Construct validity in psychological tests», *Psychological Bulletin*, 52, julio 1955, 281-302. También, American Psychological Association, American Educational Research Association y National Council on Measurements Used in Education, «Technical recommendations for psychological tests and diagnostic techniques», *Suplemento del Psychological Bulletin*, 51, 2, tomo 2, marzo 1954, y American Educational Research Association y National Council on Measurements Used in Education, *Technical Recommendations for Achievement Tests*, Washington D.C.: National Education Association, 1955.
- (3) Ministry of Education, *15 to 18: A Report of the Central Advisory Council for Education (England)*, vol. 1, Londres: H.M.S.O., 1959, p. 280.
- (4) Cook, W. W., «The functions of measurement in the facilitation of learning», en Lindquist, p. 40, *op. cit.*, nota 2.
- (5) Broudy, H. S.; Smith, B. O., y Burnett, J. R., *Democracy and Excellence in American Secondary Education*, Chicago: Rand McNally & Company, 1964, p. 88.
- (6) Cronbach, p. 284, *op. cit.*, nota 1.

VII LA EVALUACION



En este capítulo vas a encontrar:

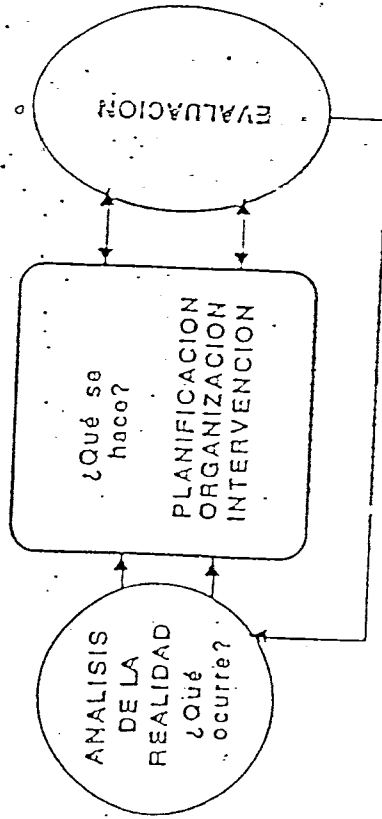
- una definición del concepto de evaluación en animación sociocultural
- una idea de cómo situar la evaluación en el proceso metodológico
- un comentario crítico de lo que hoy se hace en este campo
- criterios con los que hacer un diseño de evaluación
- algunas técnicas aplicables a determinadas situaciones

CEVERANDS, Fernando, David H. Montesinos y María Busieto. "La evaluación" en: La animación sociocultural

1. CONCEPTO

La evaluación es una parte importante del proceso metodológico de la animación sociocultural; previo al desarrollo de esta idea — a lo largo de este capítulo — se puede enunciar un concepto de evaluación que describa su función en aquel campo. La evaluación ofrece muchas posibilidades de utilización, tanto como fines a los que se puede aplicar. Es posible evaluar desde cómo se ha sentido un grupo en una sesión de alfabetización, hasta cuál ha sido la relación coste-beneficio en una agencia de publicidad.

En general, se puede decir que la evaluación significa recoger y analizar sistemáticamente una información que nos permita determinar el valor y/o mérito de lo que se hace. Determinar el valor o mérito de la intervención no es gratuito; se hace para facilitar la toma de decisiones y con el fin de aplicar lo aprendido con la evaluación a la mejora del propio proceso de intervención.



(S. 7.1. La evaluación en el proceso metodológico)

2. REFERENCIAS DESDE LA PRACTICA

Antes de pasar a desarrollar la propuesta metodológica, conviene hacer un repaso de los vicios que normalmente rodean la práctica de la evaluación en programas de tipo sociocultural. Veamos algunos de ellos:

- *La subjetividad inoperante o cuando evaluar equivale exclusivamente a valorar*
Muchas veces las evaluaciones se quedan en una simple valoración sin llegar a profundizar en los porqués y sin proponer alternativas. Esto lleva a análisis apasionados, poco operativos — «todo es un desastre» — que suelen darse con frecuencia en algunos colectivos provocando desilusión y bloqueo en los mismos. (El profesor no ha servido para nada y el grupo no participa en el curso). La evaluación conlleva necesariamente un elemento de valoración, pero no se puede quedar ahí; tiene que intentar explicar las causas y recomendar cómo mejorar lo que se está evaluando.

- *La sacralización del número o cuando evaluar equivale exclusivamente a cuantificar*
Existe una idea errónea de que evaluar significa reducir los resultados de un programa a términos cuantificables. Parece que lo serlo va siempre expresado en número. Así se pueden ver muchas evaluaciones que consisten, exclusivamente, en la utilización de indicadores de tipo cuantitativo: número de participantes, número de actividades que se han realizado, cantidad de dinero que se ha empleado... Las más sofisticadas llegan a combinar estos elementos y ofrecen ratios sobre cuánto dinero ha costado cada participante, etc. Probablemente, una de las razones por las cuales la evaluación no ha sido una práctica muy común en programas de animación sociocultural es, precisamente, esta concepción de la evaluación como exclusivamente cuantitativa y compleja en su comprensión.

- *La unidireccionalidad o la evaluación como legitimización de unos pocos*
Se realizan evaluaciones en las que se contemplan opinión.

... , valoraciones, explicaciones y recomendaciones de sólo una parte de los implicados en el programa —normalmente los responsables—. Es importante que en la evaluación se refleje lo que tienen que decir los interesados y cada uno de los sectores implicados e interesados en el programa.

- *El valor de lo absoluto y lo relativo de este valor*

En algunas ocasiones no se tiene en cuenta el valor relativo de los criterios de evaluación a aplicar. El «sólo han venido nueve», cuando la actividad es una conferencia que se ha realizado en un pueblo de 300 habitantes, debe ser necesariamente reevaluado.

- *La diferencia en la interpretación de los criterios*

Los criterios de evaluación no suelen ser los mismos para todos los implicados en ella; esto es lógico si se tiene en cuenta que, normalmente, la escala de valores de los participantes en un programa es inevitablemente heterogénea. Por ejemplo, si se evalúa como criterio positivo la «participación» de los destinatarios, puede que para algunos de los evaluadores la participación equivalga a asistir a las actividades, mientras que para otros consistirá, además, en que los destinatarios participen en la toma de decisiones sobre la marcha del programa. La solución a estas diferencias de interpretación reside en el reconocimiento de esta diferencia y en el intento de consenso en la medida de lo posible.

- *Todo en el mismo saco o la confusión en los niveles de evaluación*

Con frecuencia, se engloban o equiparan muy distintos niveles de análisis en una misma evaluación. Es distinto evaluar la institución que lleva un programa a cabo que el programa por sí mismo. También es diferente evaluar, por ejemplo, un programa de alfabetización a evaluar lo que han aprendido los alumnos receptores. Hay que diferenciar bien los niveles y saber qué es lo que queremos analizar en cada momento.

- *La recogida de datos compulsiva*

Algunas veces se confunde evaluación con la simple recogida de datos y se desarrollan registros a rellenar sobre lo que ocurre en el programa. Estos registros, en muchos casos, no son utilizados más allá de esta propia recopilación. Aunque puedan servir como un elemento de control sobre lo que se hace, no son evaluación propiamente dicha, ya que no se da ni una valoración.

ción ni una explicación de los hechos. En algunas ocasiones, se recogen muchos datos que luego no son significativos para una evaluación. Hay que ser selectivos en la recogida de información, de manera que no se empleen energías inútiles. Es más importante que se utilice a fondo la poca información recogida, a que se tengan muchos datos que no sirvan para hacer mejorar el programa.

- *La tiranía de la investigación experimental*

En ocasiones se ha entendido la evaluación como un proceso de investigación en el que han de utilizarse necesariamente los mismos procedimientos y técnicas del método científico experimental. Es importante diferenciar la investigación de la evaluación, liberando a esta última de la rigidez y lenta capacidad de respuesta que conlleva el método experimental cuando se enfrenta a situaciones que no son de laboratorio. La evaluación en la animación sociocultural debe ser algo eminentemente práctico y aplicado, en donde el factor tiempo tenga una importancia crucial. Por lo tanto, no se rechaza la utilización de este método (no cabe la menor duda que, existiendo las condiciones apropiadas, debe ser utilizado) sino que la mayoría de las situaciones evaluativas en la dinámica sociocultural no permite su utilización. En otras palabras, puede ser más interesante sacrificar la utilización del método experimental (que en algunas ocasiones puede llevar a la obtención de datos muy precisos, si bien sobre parcelas muy restringidas del programa que se está evaluando) y obtener, sin embargo, un mayor sentido global de la evaluación. Al fin y al cabo, el sentido último de la evaluación es que resulte útil para la mejora del programa o la toma de decisiones acerca del mismo. Por lo tanto, a diferencia de lo que sucede en la investigación básica y aplicada, en la evaluación de programas normalmente no es prioritaria, aunque pueda ser deseable en algunos casos, la generalización de los resultados.

- *Cuando evaluar equivale a cubrir el expediente*

Muchas veces, se evalúa simplemente por el hecho de que hay que rendir cuentas a alguien o justificar una subvención obtenida, sin confiar ni esperar que la evaluación pueda ser útil para la mejora de lo que se está haciendo. Este tipo de evaluaciones hacen desconfiar de su utilidad como parte de una metodología comprometida con los procesos socioculturales.

3. REFERENCIAS BIBLIOGRAFICAS

La conciencia de la necesidad de la evaluación en procesos socioculturales y, sobre todo, el desarrollo de la misma son fenómenos relativamente nuevos en España. En la actualidad, se podría decir que existe ya un discurso de la necesidad de la evaluación como parte inevitablemente integrante de un proceso metodológico más amplio. Este discurso, sin embargo, ha quedado muchas veces en una mera declaración de intenciones que no se ha correspondido con una ejecución real. Esto puede ser considerado como un fenómeno extraño, si se analiza la literatura existente en nuestro país sobre evaluación, que se podría dividir en varios campos:

— Literatura relacionada con el trabajo social; es aquí donde encontramos una mayor fuente de documentos sobre metodología de procesos socioculturales, que vienen de la tradición latinoamericana (Espinoza Vergara, Ander-Egg...). También en este campo es donde se han desarrollado más sistemas de evaluación que, sin embargo, tienen más que ver con evaluaciones a gran escala de servicios sociales que con otro tipo de procesos socioculturales.

— Literatura relacionada con la evaluación psicológica y diagnóstica; este ámbito tiene poco que ver con los procesos socioculturales, ya que se trata de evaluaciones ceñidas a la casuística individual.

— Literatura relacionada con la evaluación educativa; tradicionalmente este tipo de literatura ha estado centrada en la evaluación del alumno, más que en la evaluación de programas y proyectos educativos. En otros países —especialmente EEUU— sí se ha desarrollado este último tipo de evaluación y en España empieza a aumentar la preocupación sobre ello, fomentada sobre todo por el movimiento de reforma de las enseñanzas. Empiezan a traducirse algunos materiales norteamericanos de interés y las distintas Instituciones Públicas comienzan a dar mucha importancia a la evaluación de programas educativos.

— Literatura directamente relacionada con la animación sociocultural y el desarrollo comunitario; no existe prácticamente hasta el momento ningún libro ni artículo monográfico sobre el

tema. En algunos casos se dedica un espacio —normalmente muy corto— a hablar solamente de la necesidad de la evaluación. Normalmente, o bien se teoriza —muy superficialmente— sobre el tema o bien se exponen las valoraciones hechas sobre algún programa, es decir, se habla del contenido de la evaluación pero no de la metodología que se ha utilizado para llegar a ella. Salvo alguna excepción (M. Jesús Garrido, Franch y Martinell, Espinoza) no se proponen instrumentos ni técnicas concretas de evaluación.

En los términos en los que se menciona la evaluación en los documentos se puede resumir en tres aspectos:

— Evaluación como parte del proceso metodológico de intervención sociocultural (estudio de la realidad, programación, intervención y evaluación)

— Evaluación como instrumento de maduración en los procesos de grupo

— Evaluación como medición del grado de cumplimiento de los objetivos propuestos, es decir, de los resultados esperados. Algunos autores habían entonces de rentabilidad social.

El problema, sobre todo, reside en que se desarrolla escasamente la parte más estrictamente metodológica de la evaluación y cuando se hace, lo es de manera muy general y con criterios estrechos (cumplimiento de objetivos). Creemos que esta es la respuesta lógica a la falta de tradición evaluativa en programas y servicios socioculturales en nuestro país. Sin embargo, cada vez existe más preocupación y más gente experimentando sistemas de evaluación que son útiles para los programas en los que están implicados.

4. EL PLANTTEAMIENTO

Una vez aclarado el concepto de evaluación y hecho un repaso a lo que se hace normalmente y a lo que se ha escrito sobre el tema, se puede ya exponer el planteamiento del que se parte para desarrollar la propuesta metodológica que en este libro se hace.

Recordando el concepto de evaluación: recoger y analizar sistemáticamente una información que permita determinar el valor y/o el mérito de lo que se hace, y concretando un poco más, se puede decir que la evaluación se concibe como un proceso en el cual se recoge y analiza sistemáticamente una información sobre un programa, actividad o intervención con la intención de utilizar (es decir, aplicar) ese análisis en la mejora del programa, actividad o intervención.

4.1. El sentido de la evaluación

Esta definición implica, claramente, que el sentido de la evaluación reside precisamente en que ésta sea usada, de hecho, para despejar interrogantes, mejorar la efectividad o tomar decisiones en cuanto a un programa, actividad o intervención y a lo que éstos afectan se refiere. Es decir, la evaluación ha de ser ante todo *útil y práctica*:

— *Útil*, porque la información que nos ofrezca debe servir para la mejora del programa, actividad o intervención: es decir, debe ser una información fácilmente aplicable y utilizable. No nos sirve prácticamente de nada, saber que tal Ayuntamiento, ha realizado X actividades culturales este año si lo que pretendía nuestro programa es potenciar la creación de núcleos activos de personas que dinamicen el pueblo.

— *Práctica*, porque debe estar dirigida a la acción, al igual que todos los momentos del proceso metodológico de la animación sociocultural. Para eso, debe estar basada en la realidad concreta del programa evaluado.

En definitiva, si no se utiliza la información obtenida a través de la evaluación, los esfuerzos y energía empleados no merecen la pena.

algo más sobre

los criterios de utilidad

En 1981, el Comité Conjunto de Normas para la Evaluación educativa pública en los EE.UU. una serie de normas para la evaluación de programas, proyectos y materiales educativos.

Estas normas se desarrollaron en un periodo de cinco años y dicho Comité representaba a más de doscientas personas en el mundo profesional de la evaluación.

Las normas estaban divididas en cuatro condiciones, por orden de prioridad:

- utilidad
- viabilidad
- ética
- precisión.

Este Comité consideró que, primero, una evaluación debe ser útil para las personas implicadas en el programa evaluado.

Después, debe ser viable tanto en términos políticos, como prácticos, como de coste-beneficio y debe ser llevada a cabo ética y honestamente.

Finalmente, si ya hemos comprobado que es útil, viable y ética, entonces nos ocupamos de la precisión técnica.

Estas normas, aunque especializadas en programas e intervenciones educativas, y teniendo como marco de referencia la idea de evaluadores externos al programa, pueden ser útiles en el sentido de que responden a una tradición evaluativa de muchos años.

Existe un resumen de estas normas traducido al castellano en el libro de Stulliebeam y Shinkfield, mencionado en la bibliografía.

4.2 Características necesarias

Para que la evaluación sea efectivamente práctica y útil, deben darse una serie de características en dicho proceso evolutivo.

-- *Capacidad de respuesta* según las situaciones: pueden ser muy diversas o variadas. Es importante adecuarse a cada situación concreta sin empeñarse en aplicar modelos preestablecidos o recetas aprendidas.

-- *Flexibilidad metodológica*: todos los métodos de recogida de información pueden valer. Se debe estar abierto a todas las posibilidades, ya que cada situación evaluativa es distinta y puede requerir distintos procedimientos. En cualquier caso no hay una única vía para cada evaluación; debemos saber del máximo número de métodos ya que debemos elegir el (los) más adecuado(s) a cada momento. Sería interesante que en las ciencias sociales, en general, y en la animación sociocultural, en particular, se fuesen desarrollando métodos cualitativos de recogida de información, ya que estos han sido subestimados y, por lo tanto, poco estudiados y utilizados; sin embargo, no por ello se debe rechazar de lleno la utilización de métodos cuantitativos, siempre y cuando tengan sentido y sea compatible con la situación evaluativa concreta.

-- *Temporalidad*: la evaluación y sus resultados deben planearse en el momento más adecuado para su utilización. El factor tiempo es especialmente importante; normalmente cuando nos hacemos una pregunta nos urge saber la respuesta cuanto antes, máxime cuando de la contestación depende que cambemos algo en nuestra intervención. Una evaluación puede resultar inútil si pasa mucho tiempo entre que se formulan las preguntas evaluativas y se contestan las mismas.

-- *Sensibilidad social*: entendida en un sentido amplio, quiere decir que es necesario tener en cuenta todos los sectores implicados, sus intereses y razones para querer o no querer una evaluación. La evaluación debe intentar reflejar todos los intereses e intentar linar al máximo las posibles incompatibilidades.

-- *Creatividad*: hay muchas maneras de recoger información y de resolver situaciones evaluativas. Es importante que se barajen todas las posibilidades y se escoja la más adecuada y eficiente.

-- *Continuidad*: el proceso evaluativo debe tener una continuidad, ya que los resultados de la evaluación son utilizados

para mejorar o cambiar el programa que, más tarde, se volverá a evaluar para recibir de nuevo *feedback*, y así sucesivamente.

-- *Realismo*: la evaluación debe ser ajustada al tiempo, recursos y energía de que se dispone. Muchas veces puede parecer interesante evaluar todos los aspectos de un programa, pero es posible que no se tenga tiempo, ni energía, ni recursos para hacerlo; más aún, es posible que no podamos utilizar toda la información que obtengamos de una sola vez. En cualquier caso es necesario dar prioridad a los aspectos más relevantes para nuestra tarea y a los que de verdad podamos evaluar.

-- *Participación*: este es un aspecto necesario en todos y cada uno de los momentos del proceso metodológico de animación sociocultural. Todos los implicados en el programa tienen algo que decir sobre el mismo. el proceso evolutivo debe llevarse a cabo de tal manera que se dé una participación de todos los sectores implicados.

4.3. Las razones de la evaluación

La evaluación puede ser utilizada en animación sociocultural principalmente por dos razones:

• Es una manera de recibir «feedback» sobre las intervenciones y los programas que se llevan a cabo, un modo de mejorar y progresar. Al fin y al cabo, la evaluación puede ser un proceso de aprendizaje de la propia intervención.

• Es una responsabilidad social y política, especialmente si las actividades llevadas a cabo se financian con fondos públicos, ya que el contribuyente tiene derecho a saber en qué y cómo se está empleando el dinero público.

Reconociendo que ambos roles de la evaluación son importantes, en este libro se ha optado por desarrollar únicamente el primero pues, aunque también sean técnicos los que hayan de desarrollar ese segundo rol, es una responsabilidad más directamente política y más externa al propio proceso metodológico de animación sociocultural.

La propuesta metodológica que aquí se plantea propone una evaluación que tenga como rol principal el de aprendizaje y me-

lora, que sea una evaluación formativa y participativa y que sea consecuente y esté vinculada con el proceso metodológico planteado, es decir, que sea una necesidad del propio proceso de análisis, planificación e intervención.

4.4 Objetivos de la evaluación

El fin último de la evaluación es, pues, la mejora del programa, actividad o intervención. Para mejorar, dependiendo de las circunstancias específicas y de las características del programa, puede ser interesante conseguir uno o más de los siguientes objetivos:

- *Medir el grado de idoneidad, eficacia o eficiencia de un programa.*
- *Idoneidad: si se adecúa a la realidad y a lo que queremos conseguir.*
- *Eficacia: si consigue lo que se había propuesto como metas (si ha cumplido los objetivos o produce el efecto deseado)*
- *eficiencia: cuánto de bien consigue eso que se había propuesto (cuánto tiempo y recursos han sido empleados).*
- *Facilitar el proceso de toma de decisiones del colectivo y sobre la intervención, actividad o programa.*

• *Fomentar un análisis prospectivo sobre cuáles y cómo deberían ser las intervenciones futuras. Esto a su vez contribuye, en último término, al aprendizaje y conocimiento progresivo y activo, individual sobre modelos de intervención y metodología en la animación sociocultural.*

5. LA PROPUESTA METODOLÓGICA

Hasta ahora se han establecido las bases de un planteamiento metodológico: se ha hablado de la idea general de este planteamiento (el sentido de la evaluación), de las características necesarias de una buena evaluación y de las razones (los fines) y los objetivos de la evaluación en animación sociocultural.

A continuación se va a concretar más esta propuesta metodológica analizando los diferentes tipos de evaluación posibles y así como estableciendo unas bases que puedan resultar útiles para llevar a cabo un buen diseño de evaluación.

5.1 Tipos de evaluación

- *Según el papel que cumple la evaluación*

La evaluación puede cumplir dos papeles: el *formativo* y el *sumativo*.

Una *evaluación formativa* hace un seguimiento del programa o actividad. Se puede considerar que esta evaluación es una parte integral del proceso de planificación y desarrollo de un programa, y su función es ofrecer un «feedback» continuo que contribuya a la reevaluación y el aprendizaje sobre la intervención, mientras ésta se está llevando a cabo. En general, la evaluación formativa cumple la función de ayudar a los implicados en el programa a mejorar y ajustar lo que se está haciendo.

La *evaluación de tipo sumativo* determina hasta qué punto los objetivos se cumplen o se producen determinados efectos no previstos; la evaluación se realiza una vez finalizado el programa. Valora el producto final, va más encaminada a la toma de decisiones y sirve como medio de controlar responsabilidades. Sirve, a su vez, para obtener una información que permita comparar los costos, el mérito y la valía de otros programas similares.

Como se ha mencionado más arriba, esta propuesta metodológica se va a centrar más en una evaluación de tipo *formativo*. Esto, además, es coherente con la idea de que la animación sociocultural pretende fomentar procesos de participación.

• Según el contenido de la evaluación

Según el contenido de la evaluación, se puedan diferenciar cuatro tipos: evaluación de necesidades, de diseño, de proceso y de resultados¹.

Evaluación de necesidades

Esta evaluación se hace después de haber estudiado la realidad y es previa a la formulación del proyecto. Se evalúa el contexto y la realidad sobre la que se quiere intervenir y se realiza un diagnóstico de las necesidades de y con el grupo destinatario, para dirigir nuestra acción. Este tipo de evaluación coincide con el análisis de la realidad, que ya ha sido tratado previamente.

Esta evaluación sirve, sobre todo, para ayudar a desarrollar objetivos y para tomar decisiones sobre la planificación de la intervención. A su vez, cumple la función de ofrecer una base que sirva para juzgar los resultados del programa.

Evaluación del diseño

Esta es una evaluación del diseño del programa, de su coherencia y su aplicabilidad. Puede cumplir una función de ayuda, a la hora de dar forma a propuestas y proyectos y formar decisiones sobre la estructuración de los mismos. Permite, además, obtener una base para juzgar cómo se está llevando a cabo el programa.

Las bases para este tipo de evaluación se pueden encontrar en los procesos de planificación y elaboración de proyectos también tratados anteriormente.

Evaluación del proceso y desarrollo del programa

Este tipo de evaluación sirve para guiar el proceso de ejecución del programa, de manera que se obtenga una información útil para realizar los ajustes convenientes mientras el programa se está llevando a cabo. Busca, pues, explicaciones de lo que pasa, de los fallos y los cambios que se producen.

¹ En 1969, Stufflebeam propone un modelo comprensivo de evaluación al que denomina Clif, conjunto de siglas para denominar estos cuatro tipos de evaluación en función de contexto —necesidades—, del input —diseño—, de proceso y del producto —resultados

Se puede realizar a través de distintos puntos de control y llene una función de retroalimentación, de ofrecer información que permita mejorar y refinar el diseño y la ejecución del programa. Esta información, además, puede ser útil posteriormente cuando se evalúen los resultados. Así, este tipo de evaluación de proceso pretende ir ajustándose mejor a la realidad y a los destinatarios a los que se dirige.

Evaluación de resultados

Describe y juzga los resultados de un programa de intervención, relacionándolos con los objetivos y las necesidades, para evaluar el mérito y valía del programa en su conjunto. Así, sirve para evaluar los efectos que ha producido el programa, no sólo los pretendidos y positivos sino también los no buscados y negativos.

Se evalúa una vez finalizada la acción y facilita la toma de decisiones sobre la continuación, terminación o modificación del programa.

Los cuatro tipos de evaluación pueden ser importantes tanto desde el punto de vista formativo como desde el sumativo. En la tabla siguiente se relacionan entre sí distintos tipos de evaluación.

	Evaluación de necesidades	Evaluación de diseño	Evaluación de proceso	Evaluación de producto
Evaluación formativa (orientada a la toma de decisiones)	Guía para la elección de objetivos y el establecimiento de prioridades.	Guía para la elección de estrategias y procedimientos.	Guía para la ejecución.	Guía para la terminación, continuación o modificación.
Evaluación sumativa (orientada a la asignación de responsabilidades)	Registro de objetivos y de las bases para su elección.	Registros de estrategias y procedimientos y bases para su elección.	Registro del proceso real.	Registro de lo conseguido y de las decisiones tomadas.

Fig. 7.3. Tipos de evaluación

(De Stufflebeam y Shrinkfield, *Evaluación sistemática*)

5.2. Diseño de la evaluación

La evaluación es un proceso lógico y racional, cuya realización conlleva una serie de tareas de secuencia más o menos flexible. Los pasos a tener en cuenta a la hora de llevar una evaluación a cabo son:

1. *Delimitación y concimiento de lo que queremos evaluar.* Si estamos haciendo una evaluación de tipo formativo y de proceso, seguramente las personas que la llevan a cabo, sean los integrantes del equipo base, con lo que poseerán un conocimiento amplio del programa. Adn así, es necesario poner en común no sólo lo que se evalúa sino también los objetivos de la evaluación (en definitiva, conviene preguntarse de antemano para qué se va a utilizar la evaluación).

2. *Delimitación de las preguntas que queremos contestar con la evaluación (definir cuáles son los aspectos que interesa evaluar).* Es el paso más difícil, pero hay que delimitarlas bien e intentar especificarlas al máximo. Por ejemplo, no basta con formular una pregunta como: ¿Qué grado de participación ha habido? Es necesario especificar y poner en común lo que se entiende por participación en ese contexto o intervención concreta. Siguiendo este ejemplo y tomando como referente un curso de animación sociocultural, este grado de participación se podría especificar en:

- asistencia,
- intervenciones en la discusión que sean aportaciones útiles y positivas.

- predisposición a participar activamente en las actividades y tareas propuestas por el coordinador o por el grupo (actitudes positivas o neutrales o actitudes negativas pero reconocidas y proponiendo alternativas).

- planteamiento o propuesta de tareas, actividades y alternativas grupales (teniendo en cuenta al grupo y adquiriendo un compromiso posterior a ese planteamiento).

El éxito de una evaluación depende principalmente de estas tareas. Hay que establecer unas prioridades, ya que es muy posible que no podamos contestar a todo. ¿Qué es lo que más nos interesa saber? Para esta tarea de establecer prioridades de los aspectos interesantes a evaluar, se deben tener en cuenta los siguientes dos factores:

Los dos factores son propuestos por Greenbach (1982)

- Grado de *incertidumbre* sobre la pregunta que nos hacemos. ¿Cuánto sabemos sobre esa pregunta? Es posible que en algunos casos ya sepamos la contestación y no merezca la pena que se empleen esfuerzos en contestar algo que ya se sabe.
- Grado de *influencia*: se define como el peso (importante) que tiene ese aspecto o pregunta en el proceso de toma de decisiones que afectan al programa.

Estos dos factores se pueden combinar en un cuadro de doble entrada que puede orientar a la hora de establecer dichas prioridades:

	BAJA INCERTIDUMBRE	ALTA INCERTIDUMBRE
BAJA INFLUENCIA	Sólo mantener canales abiertos para obtener información accesoria 4º	Se deben emplear esfuerzos 2º
ALTA INFLUENCIA	Se debe recoger información accesoria 3º	Máxima prioridad. Se deben emplear esfuerzos 1º

Fig. 7.4 prioridades en la evaluación

Los números de cada casilla indican la prioridad de cada una de ellas.

3. *Delimitación de la evidencia o información que ya tenemos y nos puede servir para contestar a nuestras preguntas.*

4. *Decisión acerca de cómo vamos a obtener la información o evidencia que nos falta.*

5. *Diseño de instrumentos que nos permitan obtener esa información.* Sesiones de grupo, grupos de discusión, cuestionarios, entrevistas, observación, recopilación documental, etc.

6. *Recopilación de información y datos.*

7. Análisis de esos datos e información.

8. *Informe de evaluación*: es una manera de ordenar y tener presente todo lo que hemos ido descubriendo. Se deben separar los resultados de la interpretación de los mismos, así como de las recomendaciones (de esta manera, el lector puede observar cuál es el proceso lógico de la interpretación, y si no está de acuerdo, hacer su propia interpretación de los resultados). La presentación de este informe puede ser escrita o hablada. Lo importante es que los resultados de la evaluación lleguen a todos los implicados en el programa, para que se propicie al máximo la utilización de los mismos y la evaluación sea un esfuerzo que haya merecido la pena. Una idea es realizar distintos informes en función de los diferentes sectores implicados (se puede poner más énfasis en lo que más interesa a cada uno).

5.3. Criterios. ¿Qué se debe evaluar?

Cuando hablamos de criterios de la evaluación en animación sociocultural, estamos hablando de cuáles deben ser las preguntas que nos debemos formular a la hora de evaluar un programa o intervención.

Lo primero a tener en cuenta, es saber qué es lo que queremos evaluar en términos generales. Ya hemos mencionado antes que la confusión de criterios es uno de los vicios de la práctica evaluativa. Tenemos que determinar a priori si lo que se quiere evaluar son los destinatarios del programa, sus necesidades, el diseño del programa, el desarrollo del mismo o los efectos que produce.

Aunque cada programa, actividad o intervención tiene sus propias características, se puede generalizar y hacer un listado de los aspectos generales —criterios— (que habría que especificar según cada caso) más importantes, siendo conscientes de que no es una lista cerrada.

Evaluación de necesidades

- ¿Quiénes participan en el análisis de la realidad?
- ¿Cuál es la percepción social y la explicación de la realidad por parte de los posibles implicados en el programa?
- ¿Hay diferencia entre el discurso aparente y el discurso latente sobre la realidad estudiada?

algo más sobre

el modelo CIPP

Stufflebeam, siguiendo su modelo CIPP (modelo comprensivo de evaluación que incluye evaluación de Contexto, de Input, de Proceso y de Producto), habla en general de cuáles son los requisitos generales (lo que hay que evaluar) en cada tipo de evaluación.

- Contexto: necesidades, problemas, oportunidades y objetivos propuestos.
- Input: planes o programas alternativos, actividades del equipo base, calendario, equipamientos, recursos y presupuestos.
- Proceso: ejecución del proyecto o programa, nivel de participación de los destinatarios, grado de satisfacción.
- Producto: efectos y resultados que el programa ha producido, pretendidos y no buscados, positivos y negativos, inmediatos y a largo plazo.

fig.7.5.

Evaluación del diseño

- ¿Existe una coherencia entre análisis de la realidad, objetivos generales, objetivos específicos, actividades y metodologías?
- ¿Responden los objetivos a las necesidades detectadas?
- ¿Son los objetivos congruentes con la «filosofía» de la organización o grupo que los formula?
- ¿Están los objetivos específicos bien formulados con respecto a los objetivos más generales?
- ¿Están los objetivos correctamente priorizados? ¿Responden esas prioridades a las necesidades reales?
- ¿Se corresponden las actividades a la jerarquización de objetivos?
- ¿El planteamiento de la metodología es el adecuado para los objetivos y las actividades?
- ¿Existen o se prevén los recursos necesarios?

Evaluación del proceso y desarrollo del programa

En cuanto al programa en general:

- ¿Se ejecuta el programa o actividad como se pretendía?
- ¿Cuáles son los puntos fuertes y los débiles de la actividad o programa?
- ¿Puede haber una mejora con las mismas condiciones y recursos? ¿Cuáles son los ajustes necesarios?
- ¿Qué factores influyen para que el programa se esté desarrollando de esta manera?

Con respecto a los participantes:

- ¿Cómo llegan los participantes al programa?
- ¿Se llega al grupo de personas previsto?
- ¿Cómo evoluciona el grupo?
- ¿Se benefician todos los destinatarios de igual manera?
- ¿Qué grado de participación se consigue?
- ¿Qué grado de satisfacción se da?

Con respecto al contenido/actividades y metodología

- ¿La importancia relativa concedida a cada actividad o sesión es la apropiada?
- ¿Los métodos elegidos son empleados con la máxima eficacia?

En cuanto a los recursos:

- ¿Tenemos los recursos necesarios?
- ¿Se utilizan los recursos —económicos, materiales, humanos— de una manera adecuada? ¿Se pueden explotar u optimizar mejor?

En cuanto al equipo base:

- ¿Funciona el equipo adecuadamente? ¿Cómo se funciona en las reuniones? ¿Son productivas?
- ¿Están bien organizadas y distribuidas las áreas?
- ¿Asume cada uno su responsabilidad?
- ¿Existen conflictos? ¿Se resuelven de manera constructiva?
- ¿Cómo se interacciona con los participantes?

Evaluación de resultados:

- ¿Qué efectos positivos y negativos ha producido el programa?

— ¿Se han cubierto las necesidades reales de la población a la que va dirigido?

- ¿Hasta qué punto se han cumplido los objetivos?
- ¿Ha habido resultados inesperados?
- ¿Cuál es la relación entre los efectos a corto y largo plazo?
- ¿Se llegó al grupo de participantes previsto?
- ¿Se podría haber conseguido lo mismo por menos dinero, esfuerzo o tiempo, etc.?
- ¿Cuál es la valía de los costes (relación coste-beneficio)?
- ¿Funcionó el equipo base adecuadamente?

Es importante recalcar que es necesario preguntarse el por qué de las respuestas que obtenemos de cada pregunta. Esto es especialmente crucial en la evaluación del proceso, ya que es el paso que nos va a facilitar el planteamiento de alternativas o ajustes necesarios mientras el programa se está llevando a cabo.

5.4. Indicadores

Unido al concepto de evaluación se encuentra el concepto de indicadores. Este concepto ha sido utilizado, principalmente, en la evaluación de servicios sociales y se ha relacionado, casi siempre, con una evaluación de tipo cuantitativo.

Un indicador es una unidad de información que nos señala si un cierto criterio de evaluación se ha dado o no. Por ejemplo, algunos indicadores del grado de participación en un curso de intervenciones de tipo cuantitativo, «ha habido propuestas y «nº de trucillas para el grupo», «hay predisposición a participar activamente en las actividades que se proponen», «todos intervienen en la discusión» de tipo cualitativo.

Muchas veces estos indicadores, cuantitativos o cualitativos, pueden ser útiles. Sin embargo, no se puede hacer una evaluación seria de un programa de animación sociocultural sólo a base de indicadores cuantitativos, ya que perderíamos una información muy valiosa —probablemente, la que más nos interesa o resulta más significativa— que no puede ser reducida a términos numéricos.

En definitiva, formular indicadores puede ayudar a especificar los criterios o preguntas de evaluación; también a sistematizar la búsqueda de información. Sin embargo, se tendrá que contemplar otra información significativa que no se haya incluido.

do en los indicadores. Es por esta razón por la que en algunas situaciones evaluativas no conviene determinar indicadores previamente, ya que puede interesar afrontar la evaluación de la manera menos predeterminada posible (por ej: programas nuevos que se estén experimentando).

5.5. Instrumentos de evaluación. ¿Cómo se evalúa?

La evaluación consiste en una *recogida sistemática de información* focalizada en una serie de aspectos; posteriormente analizada y juzgada. Los instrumentos de evaluación pueden ser cualquier instrumento de recogida de información que se ajuste al tipo de información que queremos, a las circunstancias específicas de cada situación evaluativa y a los recursos y al tiempo disponible.

La idea de que la evaluación debe ser útil conecta con la necesidad de una flexibilidad metodológica; es decir, cada situación puede exigir o permitir la utilización de instrumentos o métodos muy distintos, y cualquiera puede ser válido, siempre y cuando nos dé una buena información para contestar los interrogantes planteados. La utilidad de una evaluación va a depender en gran medida de que los métodos utilizados sean los apropiados.

Los elementos a tener en cuenta en la elección de instrumentos de recogida de información son los siguientes:

- *Naturaleza de la información* que se quiere obtener: es distinto saber qué es lo que opina, siente o espera la gente, que lo que sabe o cómo se comporta. También es diferente si se quiere saber cómo siente o qué opina cada una de las personas (en ese caso emplearíamos métodos cualitativos individuales como la entrevista), si nos interesa más cuántos de los participantes opinan algo en concreto o se comportan de una manera determinada (métodos cuantitativos anónimos como la encuesta o la observación estructurada), o si nos interesa cómo se comporta u opina un grupo (métodos como la discusión en grupo o la observación participante).

- *Razón por la que se hace la evaluación y destinatarios de la misma:* si la evaluación va a ser sólo utilizada por las personas

UNA LISTA DE POSIBLES INDICADORES CUALITATIVOS (*)

1. Se ha solucionado el problema
2. Los participantes proponen seguir con la actividad
3. Se ha incrementado la participación de otros afines
4. Ha servido como punto de partida para la creación de un grupo
5. Ha dado resonancia a un colectivo olvidado o excluido
6. Se ha cubierto un campo sin trabajar hasta el momento
7. Se ha incrementado la comunicación entre los distintos colectivos del pueblo
8. Se ha realizado en colaboración con distintas instituciones
9. Se ha realizado en colaboración con distintos colectivos ciudadanos
10. Se ha conseguido la participación de un colectivo especialmente marginado o desfavorecido de la actividad comunitaria
11. Se ha conseguido manifestar una problemática hasta el momento latente
12. Se ha conseguido una mayor cohesión en el Equipo Base de intervención por parte de algún colectivo
13. Se ha producido economía de recursos con los mismos resultados
14. Se ha conseguido un conocimiento más aproximado de la realidad
15. Ha servido para el replanteamiento de la actividad o programa
16. Ha tenido repercusión a nivel social, en los medios de comunicación
17. Se ha conseguido eliminar o disminuir una determinada situación de marginación concreta
18. Se ha elaborado un dossier o material de documentación

(*) Lista propuesta por el Instituto de Asistencia y Servicios Sociales de la Diputación de Valencia

19. 7.6.

más directamente implicadas en el programa (en este caso hay información o datos que se pueden obviar), es diferente a si la evaluación tiene que ser fácilmente interpretable y comprensible para alguien ajeno al programa. Por ejemplo, si se tiene que rendir cuentas a un político, es muy probable que haya que ofrecerle algunos datos cuantitativos, ya que estos van a ser los más útiles para ofrecer resultados y responder políticamente ante los organismos correspondientes.

• *Recursos humanos*

que tener en cuenta desde cuándo gente disponible hay para observar o entrevistar, etc., hasta si tenemos grabadora para el grupo de discusión o disponemos de dinero para fotocopiar las encuestas.

• **Tiempo disponible:** hay que saber cuánto tiempo se dispone, sin correr el riesgo de que la evaluación sea menos útil porque esté desfasada (por ejemplo, para evaluar las actividades de un centro de formación, hay que evaluar y obtener información cuando ya esté avanzado un curso, pero antes de que se programe el siguiente para introducir los cambios oportunos). Es posible, por ejemplo, que no nos dé tiempo a entrevistar a toda la gente interesada y haya que sustituirlo por un cuestionario y discusiones en grupo.

• **Momento en el que se encuentra el programa, actividad o grupo que se va a evaluar:** por ejemplo, una discusión en grupo nos va a dar resultados muy diferentes si el grupo al que vamos a evaluar está recién formado o si llevan ya tiempo funcionando.

Es probable que al evaluar un programa o intervención utilicemos más de un instrumento de recogida de información. En cualquier caso, es también una manera de validar los datos obtenidos o de complementarlos. Por ejemplo, si se quiere saber el grado de satisfacción de los participantes en el programa, es probable que interese hacer alguna discusión en grupo, alguna entrevista individual en profundidad y, si los participantes son muchos e interesa saber qué es lo que piensan todos sobre una serie de aspectos determinados, nos convendrá hacer una pequeña encuesta colectiva. A veces lo que va a dar claves para saber qué está ocurriendo y por qué ocurre, será la contradicción entre informaciones obtenidas con distintos instrumentos. En el apartado de técnicas, proponemos algunas que implican una(s) manera(s) de recoger y organizar la información.

En la tabla 7.7 se recogen algunos de los instrumentos básicos de recogida de información y algunas de sus aplicaciones.

ALGUNOS INSTRUMENTOS BÁSICOS DE RECOGIDA DE INFORMACION

TIPO DE INSTRUMENTO DE RECOGIDA DE INFORMACION	APLICACIONES
<ul style="list-style-type: none"> • RECOPILACION DOCUMENTAL: Se recaba información de documentos escritos. • OBSERVACION: <ul style="list-style-type: none"> - Estructurada: se definen de antemano categorías que se van a observar. - No estructurada: no se definen de antemano categorías fijas a observar. (participante) El observador es un participante más. • ENTREVISTA: <ul style="list-style-type: none"> - Estructurada: cada uno de los entrevistados lleva un mismo guión fijo con preguntas concretas para hacer. - Semiestructurada o no estructurada: se lleva un guión flexible de temas a tratar o no se lleva guión. • ENCUESTA O CUESTIONARIO: Permite recabar información sin la presencia de un experto y de mucha gente simultáneamente. <ul style="list-style-type: none"> - de preguntas cerradas: la respuesta se elige entre varias respuestas predeterminadas. - de preguntas abiertas: el encuestado contesta con sus propias palabras. • GRUPO DE DISCUSION: un grupo discute sobre un tema y luego se analiza el discurso. • OTRAS TECNICAS DE GRUPO (debates, tormenta de ideas, etc.) • ESCALAS O INSTRUMENTOS ESTANDARIZADOS: son escalas o test ya preparados, como escalas de personalidad, diferencial semántico, etc. • MEDIDAS NO INTRUSIVAS: son cualquier manera de recopilar información o medir sin que eso suponga intervenir directamente para esa recopilación (Ej: nº de libros que son prestados cada mes en la biblioteca pública). 	<p>Técnica básica que no interfiere en los procesos. En cualquier caso recomendable siempre y cuando se contraste la información con la obtención de otras técnicas.</p> <p>De interés cuando lo que se observa es algo concreto y se quiere que varios observadores focalicen su atención en una misma cosa. Interesante cuando el observador forma parte del proceso en el que los observados están implicados. Más subjetiva, más rica.</p> <p>Conveniente cuando se utilizan varios entrevistadores y nos interesa saber lo que piensan el conjunto de los entrevistados sobre unos temas determinados. Oportuna cuando se sospecha que se puede obtener información interesante que no se tiene predeterminada.</p> <p>Más fácil de codificar cuando tenemos muchas respuestas. Buena para obtener porcentajes.</p> <p>Genera más información y permite que la gente se exprese en sus propios términos.</p> <p>Cuando nos interesa saber el discurso social latente.</p> <p>Adaptables a cada situación. Fomentan la participación.</p> <p>Hay pocos elaborados que sirvan para la evaluación de procesos de animación. Hay que utilizarlos con cautela, sabiendo bien los límites que tienen y lo que significan sus resultados.</p> <p>Pueden ser siempre un apoyo a otras técnicas y normalmente no suponen un esfuerzo grande adicional. Suelen ser un buen recurso poco empleado por falta de imaginación.</p>

Fig. 7.7.

7000

5.6. Los protagonistas de la evaluación ¿Quién debe evaluar?

La responsabilidad última de una evaluación de tipo formativo y de proceso (así como de la de necesidades y de diseño), recae principalmente en el equipo base. Cuando se habla de responsable de que se evalúe el programa es el que está trabajando y de que se implique, en esa evaluación, a las personas oportunos. Esto no debe ser incompatible con el hecho de que la información básica recogida provenga, principalmente, de los destinatarios del programa. Como ya se ha dicho antes, el proceso evaluativo ha de ser participativo en todo momento; los datos de la evaluación han de estar disponibles, para que todos los implicados en el programa puedan efectuar su propio análisis. Es más, una evaluación de tipo formativo cumple una función de aprendizaje y *feedback* no sólo para el equipo base, sino también para los participantes en el programa.

En determinadas ocasiones, puede resultar interesante hacer evaluaciones de conjunto; esto requiere un esfuerzo y un tiempo adicional a lo habitual en la dinámica de un programa o servicio; en estos casos convendrá contar con la presencia de un evaluador o equipo evaluador externo. Aun así, el protagonista del equipo base y de los participantes debe continuar asegurado.

5.7. Los tiempos de la evaluación

Como ya hemos dicho antes, la evaluación es una parte del proceso de planificación y debe ser *planificada* en sí misma, siempre considerándolo como un proceso flexible que se adapte a la realidad concreta que se evalúa.

Dependiendo de las preguntas y la información que se desea obtener, se intervendrá en un momento u otro:

1. Antes de comenzar la actividad

Haciendo una evaluación de necesidades y una evaluación del diseño, antes de poner el programa en funcionamiento.

2. Durante la actividad

Se puede contestar a diferentes preguntas durante la actividad para ir corrigiendo y perfeccionando la actuación sobre la marcha. Esto se adecúa al rol formativo de la evaluación.

3. Inmediatamente después de la actividad

La valoración se hace cuando la actividad ha terminado incorporando la información y los datos que se recogieron, mientras la actividad estaba en marcha; este solapamiento de tiempos, refuerza la conveniencia de planificar la evaluación con antelación.

Este tipo de evaluación se adecúa al rol sumativo de la evaluación, aunque puede también tener un carácter formativo, si la actividad se repite sistemáticamente (por ejemplo: curso de monitores de tiempo libre, campamentos organizados por una misma institución o grupo, etc.).

4. Cierta tiempo después de que la actividad haya sido llevada a cabo

Se refiere a lo que normalmente se denomina *seguimiento*; dependiendo de la actividad, puede interesar conocer cuáles son los efectos a largo plazo y su diferencia con los efectos a corto. Este tipo de evaluación requiere tener cierta información del momento en el que la actividad se llevó a cabo.

Aquí se ofrecen algunas técnicas que sirven, sobre todo, para evaluar procesos en los que un grupo es protagonista. En cualquier caso pueden ser guías o maneras de organizar la información de cualquier otra evaluación:

6.1. Valoración-Explicación-Aplicación

Ya se mencionó al principio del capítulo uno de los vicios de los procesos de evaluación: equiparar estos a una simple valoración de los hechos, sin preguntarse el por qué ni presentar alternativas. Esto es contrario al concepto de evaluación que aquí se ha venido desarrollando, que implica necesariamente la idea de evaluar para mejorar, es decir, cambiar cosas para progresar o ajustarse a las nuevas necesidades.

Esto es especialmente importante cuando lo que se evalúa son procesos de grupo. Los grupos, al hacer una evaluación de sí mismos o de la tarea en la que están inmersos, tienden a quedarse en la simple valoración, normalmente egocéntrica. Cuando se evalúan procesos, es crucial preguntarse este «por qué», de las valoraciones, ya que es lo que nos va a ayudar a plantear alternativas y a poder aplicar la evaluación.

Una forma de recordar estos pasos es el siguiente cuadro, que sirve para organizar la información obtenida durante el proceso evaluativo y extraer conclusiones. Es útil también para evaluar en una sesión de grupo, a través de una discusión, y sacar partido de situaciones evaluativas grupales.

Los pasos para utilizar este cuadro, serían los siguientes:

1. Definición de criterios o preguntas evaluativas. De aquí se obtendrán las dimensiones a evaluar.
2. Valoración de cada una de las dimensiones a evaluar.
3. Explicación de las valoraciones que se hacen. Contestar al por qué de cada valoración.



FICHA DEL PROCESO DE EVALUACION

T.7.1

Dimensión a evaluar	Valoración ¿Qué?	Explicación ¿Por qué?	Aplicación Alternativas

5.6. Los protagonistas de la evaluación ¿Quién debe evaluar?

La responsabilidad última de una evaluación de tipo formativo y de proceso (así como de la de necesidades y de diseño), recae principalmente en el equipo base. Cuando se habla de responsabilidad última, se hace referencia a que el equipo es el responsable de que se evalúe el programa en el que está trabajando y de que se implique, en esa evaluación, a las personas oportunos. Esto no debe ser incompatible con el hecho de que la información básica recogida provenga, principalmente, de los destinatarios del programa. Como ya se ha dicho antes, el proceso evaluativo ha de ser participativo en todo momento: los datos de la evaluación han de estar disponibles, para que todos los implicados en el programa puedan efectuar su propio análisis. Es más, una evaluación de tipo formativo cumple una función de aprendizaje y feedback no sólo para el equipo base, sino también para los participantes en el programa.

En determinadas ocasiones, puede resultar interesante hacer evaluaciones de conjunto; esto requiere un esfuerzo y un tiempo adicional a lo habitual en la dinámica de un programa o servicio; en estos casos convendrá contar con la presencia de un evaluador o equipo evaluador externo. Aun así, el protagonista del equipo base y de los participantes debe continuar asegurado.

5.7. Los tiempos de la evaluación

Como ya hemos dicho antes, la evaluación es una parte del proceso de planificación y debe ser planificada en sí misma, siempre considerándolo como un proceso flexible que se adapte a la realidad concreta que se evalúa.

Dependiendo de las preguntas y la información que se desea obtener, se intervendrá en un momento u otro:

1. Antes de comenzar la actividad

Haciendo una evaluación de necesidades y una evaluación del diseño, antes de poner el programa en funcionamiento.

2. Durante la actividad

Se puede contestar a diferentes preguntas durante la actividad para ir corrigiendo y perfeccionando la actuación sobre la marcha. Esto se adecúa al rol formativo de la evaluación.

3. Inmediatamente después de la actividad

La valoración se hace cuando la actividad ha terminado incorporando la información y los datos que se recogieron, mientras la actividad estaba en marcha; este solapamiento de tiempos, refuerza la conveniencia de planificar la evaluación con antelación.

Este tipo de evaluación se adecúa al rol sumativo de la evaluación, aunque puede también tener un carácter formativo, si la actividad se repite sistemáticamente (por ejemplo: curso de monitores de tiempo libre, campamentos organizados por una misma institución o grupo, etc.).

4. Cierta tiempo después de que la actividad haya sido llevada a cabo

Se refiere a lo que normalmente se denomina "seguimiento"; dependiendo de la actividad, puede interesar conocer cuáles son los efectos a largo plazo y su diferencia con los efectos a corto. Este tipo de evaluación requiere tener cierta información del momento en el que la actividad se llevó a cabo.

TÉCNICAS APLICADAS A LA EVALUACIÓN

EL JUICIO

T.7.2

Pasos a seguir:

- 1.- PROPUESTA de los problemas o criterios a enjuiciar.
- 2.- SELECCIÓN de los problemas (se establecen prioridades) y de las partes opuestas así como el equipo de jueces.
- 3.- PREPARACIÓN de los argumentos formales:
 - se desarrollan los aspectos específicos del enfrentamiento.
 - Cada equipo estudia los criterios a evaluar.
 - se estudian documentos y datos para utilizarlos como argumentos.
- 4.- LA AUDIENCIA:
 - SESIÓN PREVIA: los dos equipos examinan sus argumentos y junto con el equipo de jueces se desarrollan las normas y procedimientos de la audiencia.
 - AUDIENCIA: puede constar de más de una sesión. Tras la presentación de los argumentos y los interrogatorios, el jurado delibera, o recomienda.

(Basado en el Modelo Judicial de Wolf en Stufflebeam y Shrinkfield, Evaluación Sistemática).

4. Presentación de alternativas o propuestas de cambio, continuidad o reforzamiento. Este último paso facilita la aplicación de los resultados de la evaluación.

Como ya se ha dicho antes, esta técnica es especialmente recomendable en sesiones grupales. Sin embargo, puede ser válida para organizar la información de cualquier tipo de evaluación; una vez definidos los criterios a evaluar se colocan en la columna de la izquierda, y el resto de las columnas se rellenan a medida que avanza el proceso evaluativo.

6.2. El juicio

Esta técnica parte de la simulación de un juicio. Este ha de prepararse con anterioridad, estableciendo dos grupos opuestos (uno a favor y otro en contra de lo que se está evaluando), así como un equipo de jueces imparciales que dirigen y coordinan la audiencia (aunque el dictamen final pueda ser realizado por toda ella).

Es posible que la gente implicada en el programa o actividad a evaluar tenga cosas tanto negativas como positivas que decir; en este caso, los que actúan en la parte fiscal van a tener que argumentar en lo negativo y los de la parte defensora en defenderse de las acusaciones y en demostrar lo positivo. En cierto modo, se va a utilizar parte de la técnica del «role-playing».

Algunos de los procedimientos son:

- Las reglas del juicio son flexibles pero deben estar predefinidas por los directores de la audiencia y las dos partes implicadas.
- Los directores del juicio (equipo de jueces) son los que deciden si una información es relevante o no.
- Puede haber testigos a los que se les puede interrogar.
- También se puede llamar a declarar a «expertos» en la materia que se está evaluando.
- También pueden participar otros grupos interesados o implicados en el programa evaluado.
- Puede pedirse a la parte que actúa como fiscal, que informe previamente a la parte defensora de los cargos que se le imputan, teniendo derecho a admitir por anticipado ciertos cargos y a recusar otros (esto es útil para no emplear el juicio en demostrar lo que todos ya sabemos).

- Disminuye la tendenciosidad inconsciente, ya que se va-
loria lo negativo y lo positivo.
- Se promueve una gran implicación del grupo(s) de inte-
resados.
- Se profundiza en la información y en la causa de las va-
riaciones.

— No es incompatible con que se den otras actividades eva-
luativas antes o después del juicio. No debe sustituir a las planifi-
caciones de recogida y análisis de datos de una evaluación más
amplia.

- La calidad de la información es muy alta ya que todas
las afirmaciones son probadas.
- Permítele que mucha gente exprese su punto de vista.
- La información puede tener mucho alcance por el número
de personas implicadas en la actividad.

Limitaciones

- Se requiere bastante tiempo de preparación y costes al-
tos ya que se debe contar con mucha gente para su realización.
- No siempre hay diversidad de opiniones.
- El juicio en sí mismo, es decir la actuación del mismo,
se puede convertir en algo más importante que el contenido.
- Puede depender de la habilidad de cada parte contin-
ente que el lado positivo o el lado negativo sea el más atractivo.

6.3. Matrices de evaluación

Esta es otra forma de recopilar y organizar información útil
para una evaluación. Esta técnica está basada en la idea de que,
ya que toda evaluación implica necesariamente una emisión de ju-
icios sobre lo que se evalúa, éstos no deben emitirse sin tener
en cuenta la explicación de los hechos o datos que nos inducen
a juzgar de una manera u otra.

Consiste en desarrollar tres tipos de matrices en las que se
recopila y organiza otros tantos tipos distintos de datos:

- **Matriz de descripción:** en esta matriz se contemplan dos
columnas en las que se incluye la siguiente información:

- 1.ª columna: las intenciones, lo que se pretende y planifi-
ca. Aquí se debe tener en cuenta no sólo los objetivos explícitos,
sino también los implícitos, así como las variaciones que estos
hayan sufrido a lo largo del programa.
- 2.ª columna: las observaciones, lo que realmente ocurre
y se puede observar.

• **Matriz de explicación:** en esta segunda matriz, se hace un
análisis del grado de congruencia entre lo pretendido y lo obser-
vado. Siempre va a haber una diferencia entre ambos aspectos;
lo que se trata de analizar es cuál es su embergadura y los por-
qués de la misma. Si se para sólo en analizar el hecho de la dife-
rencia, es posible que se calga en el error de mirar sólo lo negativo
del programa, cuando probablemente existirán aspectos que ha-
yan resultado positivos. Incluso aunque no estuviera contempla-
do en los objetivos iniciales del programa. Esta es la razón de
que en la matriz de explicación se incluyan columnas en las que
se reliegan no sólo los errores a cambiar, sino los aciertos a po-
tenciar. Estos errores y aciertos son los que van a explicar las cau-
sas y el grado de diferencia entre las intenciones y lo observado.

- 1.ª columna: grado de diferencia entre lo pretendido y lo
observado.
- 2.ª columna: se enumeran los errores. Estos errores van
a ser las posibles causas de esa diferencia.
- 3.ª columna: se trata de dar una alternativa a esos erro-
res. Se trata de ver qué es lo que hay que cambiar.
- 4.ª columna: se enumeran los aciertos del programa, en
el sentido de que contribuyen a que la diferencia entre lo inten-
cionado y lo observado sea pequeña y, también, por lo que de
inesperado puedan tener y que conviene potenciar.
- 5.ª columna: en ésta se resalta lo que se cree convenien-
te potenciar de esos aciertos.

• **Matriz de juicio:** en esta matriz se pretende recopilar la in-
formación relacionada con la emisión de juicios sobre el progra-
ma que se está evaluando. Esta información se recoge en dos
columnas:

- 1.ª columna: se enumeran las normas que se van a utili-
zar para emitir los juicios. Las normas son los criterios explícitos
para valorar la calidad de un programa. Hay dos tipos de normas:
— **Absolutas:** vienen dadas por las convicciones personales
de los profesionales implicados en la evaluación, sobre lo que
es bueno o deseable para un programa. También pueden venir

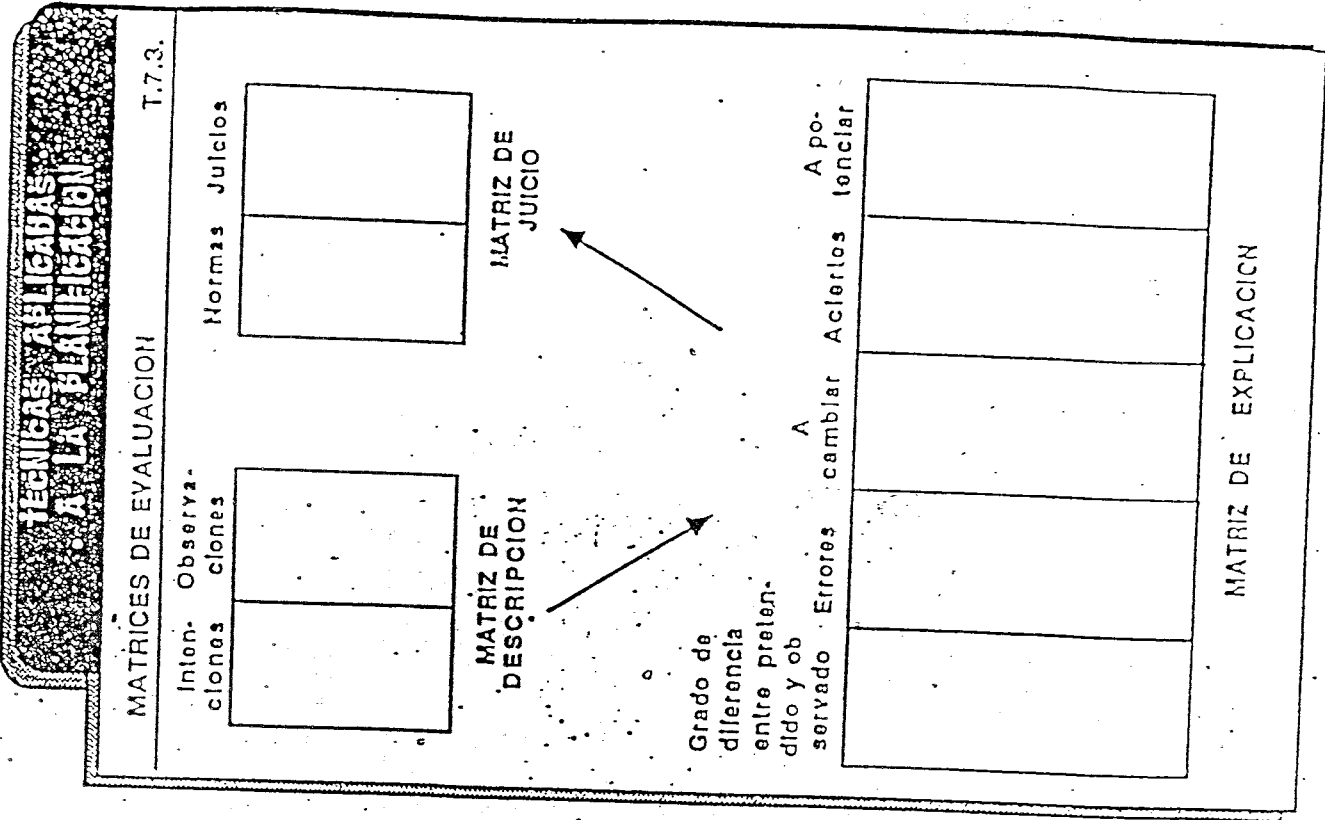
dadas por lo que se ha escrito sobre el tema — la literatura existente

— Relativas: vienen dadas por las características de programas similares que se consideran satisfactorios. En este tipo de normas también hay que tener en cuenta el análisis de la realidad en la que está inmersa el programa evaluado, así como las condiciones en las que dicho programa se ha llevado a cabo.

2.ª columna: se realizan y enumeran las valoraciones o juicios sobre el programa en general y sobre cada uno de sus aspectos, en particular.

El valor de esta técnica radica en la organización que realiza una cantidad de información considerable con respecto al programa. A su vez, y como ya se ha mencionado, no permite que se hagan las valoraciones finales sin haber contemplado antes la explicación de los hechos y la proposición de alternativas.

Su desventaja está en la complejidad que conlleva la organización de todos esos datos. Por esta razón, es una técnica más adecuada para las evaluaciones que, o bien se realizan desde el equipo base o bien desde un equipo de evaluadores externos no implicados directamente en el programa.



7. RESUMEN

La evaluación es un proceso de recogida sistemática de una información que, en su aplicación, no sólo va a ayudar a valorar sino también, y sobre todo, a mejorar lo que se está evaluando.

El criterio más importante que debe seguir una evaluación es el de utilidad, es decir, que sirva para algo y que haya seguridad de que se va a aplicar lo que se ha aprendido de ella. Para esto, la evaluación debe cumplir una serie de condiciones: —flexibilidad metodológica, capacidad de respuesta, sensibilidad social, creatividad, participación, continuidad, temporalidad, realismo — y no utilizarse como en muchas ocasiones se hace — para cubrir el expediente —, reduciéndola exclusivamente a valorar, cuantificar o recopilar datos compulsivamente, teniendo en cuenta sólo una parte de los implicados, confundiendo niveles o interpretando los criterios heterogéneamente.

La propuesta metodológica aquí expresada establece distintos tipos de evaluación y sugiere un modelo que da especial importancia a la formulación de preguntas o criterios de evaluación, y a la contestación de las mismas a través de indicadores e instrumentos de recogida de información.

Por último, se han sugerido tres técnicas distintas que pueden ser útiles para recopilar, organizar y generar información para realizar evaluaciones de procesos socioculturales.

RECOPILACIÓN DE DATOS

De todos los aspectos propios de la investigación de campo, la recopilación de datos ha sido el más tratado en la bibliografía sobre los métodos. Para no pecar de prolijo, el presente comentario sobre los métodos de datos se reducirá al mínimo. Revisaré los principales temas y cuestiones relativos a la recopilación de datos y remitiré al lector a las fuentes bibliográficas más importantes para un análisis más pormenorizado.

Un enfoque de la recopilación de datos en el campo consiste en hacerla tan intuitiva —o tan radicalmente intuitiva— como sea posible. Se cree que con una observación participativa, intensiva y de largo plazo en un contexto de campo, iniciada sin ninguna expectativa conceptual previa que pudiera limitar la ecuanimidad del investigador de campo respecto del carácter singular de la experiencia en el escenario surgirá por inducción un sentido intuitivo de las preguntas pertinentes para la investigación y de las conclusiones sobre los patrones existentes. Desde este punto de vista, el trabajo de campo se percibe como un proceso casi místico, imposible, en esencia, de ser enseñado. La mejor capacitación se obtiene a través de cursos de antropología y/o sociología que brinden una base sólida de conocimientos. Una vez que ha aprendido los fundamentos teóricos correspondientes y que ha estudiado los resultados empíricos de las investigaciones de campo realizadas, el investigador novato se traslada al campo y comienza su trabajo.

Los antropólogos, en especial, han sostenido esta concepción mística del trabajo de campo como actividad imposible de ser enseñada. Se cuenta que en cierta ocasión, un estudiante avanzado le pidió a Alfred Kroeber que lo aconsejara sobre los métodos del trabajo de campo, pues estaba por emprender un estudio sobre una comunidad norteamericana nativa en cierto lugar de California. Kroeber le hizo las siguientes recomendaciones:

1. En primer lugar, encuentre a sus indios (es decir, no vaya a estudiar otro grupo por equivocación).
2. Las libretas de apuntes y los lápices son muy útiles.
3. Asegúrese de llevar una sartén, pero no se la preste a nadie porque podrían no devolverla.

Esta es una noción muy romántica del trabajo de campo. Se llega al campo sin ninguna concepción previa y se aprenden los métodos al practicarlos (así como alguien puede aprender a nadar por haber sido arrojado al agua). Después de una tremenda tensión emocional, finalmente se inducen categorías analíticas básicas. La probabilidad de sufrir tensiones es mayor si además de experimentar un trauma emocional durante la estancia en el campo, también se contrae alguna enfermedad exótica y debilitante, como la malaria. Sólo después de haber vuelto a

... se debilita el código de los hechos de vida locales y se resuelve el
... análisis de la interacción.

... deliberativo como sea posible. Es este enfoque el que apoyamos en
... ciencia contentiva área ni en la psicología cognitiva, para la concep-
... a tomar el trabajo de campo, según la cual el investigador llega
... pagar con el agente como *tabula rasa* y por tanto tan sólo un cepillo
... dientes y un cepillo de monte. Se puede argumentar que no existen
... funciones puras, siempre aportamos a la experiencia ciertos *marcos*
... interpretación o esquemas. Desde este punto de vista, la tarea del
... de la interpretación consiste en tomar cada vez mayor conciencia de los
... roles de interpretación de los sujetos que observamos y de nuestros
... roles en los de interpretación culturalmente aprendidos, que hemos
... ido con nosotros al campo. Esto significa desarrollar una visión dis-
... ta de nosotros mismos de la línea divisoria entre observador y observado;
... lo que FRANKSON (1963, págs. 7 y 8) denominó la *visión social* este
... serena el etnógrafo

... cuando concebimos el trabajo de campo como un proceso de inda-
... fin deliberada en un contexto (véase PELTO y PELTO, 1977; LEVINE,
... ALMIRE, WEISSER y TURNER, 1980). Podemos considerar que la tarea
... recoger datos por parte del observador participante equivale a la re-
... lación progresiva de problemas, en la que se juntan aspectos relativos
... nuestro, a la formulación de hipótesis y a la verificación de hipótesis.
... cantidad clara de los investigadores de campo en el lugar la guían
... razones premeditadas sobre el nuestro y también reacciones intuiti-
... vas. Cuando y dónde acuden estos observadores, con quién hablan, a
... en observan, con quién participan más activamente en los quehace-
... cotidianos y con quién participan desde un puesto de observación
... distante, son aspectos que constituyen decisiones estratégicas acerca
... la validez de las preguntas clave para la investigación y de las hipó-
... teorizadoras del estudio.

... sin embargo, no todas las decisiones son premeditadas en la investi-
... gación. Por esta razón, la corriente del cepillo de dientes y el cuchillo
... mente tiene la ventaja de señalar la impurancia de la inducción
... perspectiva directa e intensiva en el contexto. Desde la perspectiva
... que se son más deliberativa del trabajo de campo, no obstante, la
... que no metodológica fundamental es asegurar una relación constante,
... que el investigador, entre las preguntas de investigación y la recopilación
... datos. Esto se puede lograr sin poner freno a la intuición y a la
... que el hecho de encuadrar explícitamente las preguntas de
... investigación y de buscar datos pertinentes en forma premeditada po-
... que favorezca la intuición en lugar de entorpecerla.

... que se emplea un enfoque deliberativo del trabajo de campo,
... que se emplea el análisis de los datos después de haberse retirado del
... que se emplea algunos problemas típicos de inadecuación de las
... que se emplean. Estos problemas podrían haberse evitado si se han-

... tomaron otras decisiones estratégicas en la etapa de la recopilación
... con los datos, cuando aún era posible clarificar los errores. Hay otros
... tipos principales de inadecuación de las pruebas.

1. **Volumen inadecuado de pruebas.** El investigador tiene pocas pruebas para respaldar ciertas afirmaciones fundamentales. En sus sesiones diarias, no incluyó las escenas en las que podría haber recogido las pruebas que habrían confirmado esas afirmaciones.
2. **Variedad inadecuada de tipos de pruebas.** El investigador no tiene pruebas correspondientes a un espectro de fuentes de distinto tipo (por ejemplo: observación directa, entrevistas, documentos del entorno) para respaldar afirmaciones fundamentales a través de la triangulación. No buscó triangular los datos mientras estaba en el lugar.
3. **Interpretación incorrecta de las pruebas.** El investigador no entendió correctamente los aspectos fundamentales de la complejidad de la acción o de las perspectivas de significado apoyadas por los actores en el contexto. (Su participación en escenas estratégicas reiteradas no fue suficientemente prolongada o intensiva, y/o la entrevista y la observación directa no se complementaron una a otra; y/o el investigador fue engañado por informantes que mintieron o fingieron porque no confiaron en él o no estuvieron de acuerdo con sus objetivos).
4. **Pruebas impugnativas inadecuadas.** El investigador carece de datos que pudieran impugnar una afirmación clave. Lo que es más, carece de pruebas de haber realizado una búsqueda deliberada de datos potencialmente impugnativos mientras estaba en el sitio. Esto debilita la verosimilitud de la ausencia de pruebas impugnativas y hace que el investigador pueda ser acusado de buscar sólo las pruebas que corroborarían las interpretaciones que él defende. (Al no haber establecido, mientras estaba en el campo, las relaciones entre las preguntas de investigación y la recolección de datos, el investigador no buscó pruebas que pudieran impugnar la afirmación en cuestión o que pudieran representar casos discrepantes, cuyo análisis en el campo podría haber arrojado nueva luz sobre la afirmación y sus pre-suposiciones técnicas).
5. **Análisis inadecuado de casos discrepantes.** El investigador no inspeccionó la serie de sucesos impugnativos, examinando cada uno de ellos lo sea, cada caso discrepante) y la comparó con los sucesos confirmatorios para determinar qué rasgos del caso impugnativo eran iguales y diferentes de los rasgos análogos de los casos confirmatorios. Este análisis comparativo de los rasgos suele revelar fallos en la afirmación original que si se reexaminan puede dar cuenta de los casos discrepantes, así como de los que inicialmente se creyó que eran hechos impugnatorios. Los resultados contrastantes de los hechos impugnatorios son casos, por lo tanto, que no se reconocen del conjunto de hechos impugnatorios. Los resultados contrastantes se ilustran mediante ejemplos tomados del aula en el próximo capítulo. El punto a destacar aquí es que el análisis de casos discrepantes le permite al investigador afinar y ajustar las afirmaciones, principales y sus pre-suposiciones técnicas. Si no se efectúa dicho análisis, pueden producirse dos tipos de errores: al el investigador rechaza prematuramente una afirmación, considerando como similares todos los sucesos que parecen, en un primer análisis, ser impugnativos, y/o el investigador no afirma ni afirma

Las afirmaciones que en un primer análisis parecieron estar confirmadas por los datos.)

Aspectos del acceso al lugar y de la ética de la investigación

Una investigación de campo potencialmente buena puede verse comprometida desde el principio por una negociación inadecuada del acceso al contexto. Esto origina problemas de calidad en los datos y en la ética de la investigación. El interés del investigador es tener acceso al mayor volumen de información posible. Dado los problemas potenciales de la adecuación de las pruebas arriba mencionados, la aspiración del investigador de campo es poder observar cualquier lugar del escenario en cualquier momento, y poder entrevistar a cualquier miembro de la comunidad acerca de cualquier tema. Esto puede o no coincidir con los intereses de las personas del lugar. Se suscitan problemas de intereses especiales y de riesgos especiales, no sólo entre los miembros del contexto y las autoridades externas (el personal del distrito, escolar y la comunidad local, la secretaría de educación del Estado, los organismos educativos federales), sino también dentro y entre los niveles de los sistemas en la organización (los intereses de los docentes en relación con un director, los intereses de los alumnos en relación con los docentes).

Hay que tener en cuenta dos principios básicos. Los individuos estudiados, en especial los que se estudian como sujetos centrales de la investigación, deben estar a) lo mejor informados que sea posible acerca de los propósitos y las actividades de la investigación que se llevará a cabo, así como de cualquier exigencia (trabajo, adicional) o riesgo que les pudiera acarrear el hecho de ser estudiados, y b) lo más protegidos que sea posible de cualquier riesgo. Los riesgos en cuestión pueden ser mínimos. No se trata de riesgos de índole física, como en algunos experimentos médicos, sino que el tipo de riesgo que habitualmente entraña la investigación de campo es de carácter psicológico y social (situaciones embarazosas y/o posibilidad de ser objeto de sanciones administrativas). De todos modos, puede haber un gran riesgo de ocasionar daños psicológicos y sociales cuando el trabajo de campo es efectuado por un investigador que ignora los aspectos institucionales en juego y que no ha anticipado adecuadamente los distintos tipos de daño a que están expuestas las personas de diferentes posiciones sociales en el lugar.

Los riesgos suelen ser mayores cuando atañen a las relaciones entre miembros de grupos de intereses divergentes dentro del contexto. Normalmente, el hecho de presentar informes ante un auditorio compuesto de gentiles no expone a los individuos locales a ningún riesgo. Pero el acto de suministrar informes dentro del contexto puede entrañar riesgos y esto debe tenerlo en cuenta el investigador al negociar su acceso a la información sobre los individuos del lugar.

El investigador se encuentra en una situación compleja. Debe haber acordado una etnografía del contexto a fin de anticipar los posibles ries-

gos y exigencias a que se verán expuestos los individuos estudiados. Si bien no es posible, al principio, anticipar todos los problemas éticos que habrán de plantearse, sí es posible prever muchos de ellos y efectuar las correspondientes negociaciones con aquellos grupos de interés en el lugar, cuya existencia y cuyas circunstancias son evidentes desde un principio. En los entornos escolares, estas clases generales de grupos de interés son los estudiantes, los docentes, los padres, los directores y los administradores. Algunos de los intereses divergentes de estas distintas clases de grupos, se pueden identificar de antemano. En general, es conveniente, por ejemplo, garantizar a los profesores que ciertos tipos de información sobre su práctica docente no quedarán al alcance de sus supervisores inmediatos, y que cierta información sobre la vida en el hogar de un alumno no será puesta en conocimiento de los profesores. Esta información sobre los docentes y sobre los hogares de los alumnos, sin embargo, podría no implicar riesgos para los individuos si se expusiese en forma conjunta: respecto de todos o de muchos de los docentes y respecto de todos o de muchos de los hogares.

En determinadas circunstancias, los riesgos que normalmente afectan a una posición institucional dada pueden no existir. Por ejemplo, la información de que un profesor no utiliza el libro de lectura indicado por los reglamentos del sistema escolar podría ser desfavorable para ese docente a criterio de un director de escuela, mientras que para el director, con otro punto de vista, esa misma información podría favorecer al profesor.

El investigador debe negociar los intereses particulares de los participantes especialmente vulnerables en el lugar. Las suites centrales son particularmente vulnerables, al igual que aquellos que son los únicos ocupantes de una determinada posición institucional (hay un solo director y muchos docentes, pero un solo profesor de jardín de infancia). Las personas que no son sujetos centrales del estudio probablemente corran menos riesgos. (Por ejemplo, si se están estudiando los alumnos de bajo rendimiento de sexo femenino, los varones de bajo rendimiento tienen menos posibilidades de estar expuestos a riesgos. Estos aparecerían, en los informes descriptivos, como parte del medio circundante, y no en un primer plano.) En un estudio que dirige sobre un extenso sistema escolar urbano, por ejemplo (Cazden, Carrasco, Maloof, Givens y Erickson, 1980), realizamos una observación participativa y grabaciones en vídeo, durante todo un año, de un grupo de profesores bilingües de origen hispánico que no tenían títulos oficiales. Los profesores no eran funcionarios, sino que tenían contratos renovables anualmente. En este caso, negociamos un acuerdo con el director

del establecimiento, con el superintendente del distrito y con un superintendente adjunto de la oficina escolar, por el cual nunca se les pedían a los investigadores que anotaran que anotaban en el campo a ninguna de las instituciones escolares, y estas no le solicitaran a los investigadores que les dieran su opinión sobre el profesor estudiado, ni que les declararan ver las grabaciones a fin de evaluarlo. Los investigadores tenían previsto mostrar parte de las grabaciones al director y al censo de la escuela, pero planeaban hacerlo después de ese año, una vez que tuvieran todos los datos y que los profesores hubieran revisado las grabaciones a ser mostradas y hubieran dado su consentimiento para que otros las vieran. En ningún caso se mostraría ninguna grabación a nadie fuera del grupo de investigación sin el consentimiento de los profesores. Además hicieron un acuerdo por escrito con las autoridades escolares, por el que ellas se comprometían a no solicitar que se les mostraran ni siquiera las grabaciones que los profesores habían consentido que vieran, hasta después de haber firmado el contrato anual con los docentes. Con esto no se eliminó totalmente el riesgo para los enseñantes, pero sí se redujo al mínimo el riesgo de que fueran objeto de críticas informales mientras estaban siendo observados. De hecho no hubo ninguna solicitud por parte de las autoridades para ver las grabaciones durante el año de la observación, pero de todos modos fue prudente haber previsto esa posibilidad al negociar el acceso al campo de estudio.

El investigador debe tener el cuidado de ser muy explícito acerca del uso de la información y del acceso a la misma, dado que le interesa tener el máximo acceso posible en el contexto en condiciones de suma confianza y entendimiento. El mismo proceso de negociar explícitamente el acceso en el campo con todas las categorías de personas que pueden verse afectadas por la investigación puede también crear las necesarias condiciones de confianza. En consecuencia, vemos que la responsabilidad ética y la adecuación científica van juntas en la investigación de campo. En los sujetos de la investigación consistentes libremente en que se los observe, habiendo sido informados acerca de los propósitos de la investigación y de los riesgos que esta puede acarrearles (tal como de los posibles beneficios), entonces los enseñantes y las mentiras se vuelven al mismo va que constituyen una existencia pasiva ante la presencia del investigador.

En suma, la negociación del acceso es un proceso complejo. Comienza con la primera carta o llamada telefónica al lugar. Continúa durante el transcurso de la investigación, y prosigue después de que el investigador se ha retirado, durante el proceso de análisis de los datos y durante el acceso a la investigación en condiciones convenientes, tanto para los sujetos de la investigación como para el investigador, si esta es posible. A partir de las cuales se pueden entablar relaciones francas y transparentes. Sin esas bases, la confianza mutua se hace problemática y no conseguiremos la capacidad del investigador para identificar las relaciones de significado de las personas del lugar. El tema de la ne-

gociación del acceso ha sido objeto de una considerable atención en la bibliografía referente a los métodos de la investigación de campo. Véase, sobre todo, Auer (1980, págs. 42-62), Berman, y HURMAN (1982, págs. 120, 123), SCHWARTZMAN y STRAUSS (1974, págs. 18-33) y WAN (1971, págs. 15-20, 84-93, 143-174).

El desarrollo de una relación de cooperación con los principales informantes

La confianza y la armonía en el trabajo de campo no son meramente una cuestión de buenos modales; es imprescindible que se entable una relación mutuamente satisfactoria y no coercitiva entre los informantes fundamentales y el investigador para que éste pueda hacerse una idea clara y válida del punto de vista del informante. Dado que conocer la perspectiva del informante es esencial para el buen resultado de la investigación, es necesario ganarse su confianza y conservarla durante todo el transcurso del estudio.

Una fuente de dificultades en relación con la confianza es la tendencia de los informantes a suponer, sea cual fuere el modo en que el investigador planteó los propósitos de la investigación durante las etapas iniciales de la negociación de su acceso, que dichos propósitos son, de algún modo, evaluativos. A menudo se hace necesario volver a explicarle varias veces los objetivos de la investigación al mismo informante. También es imprescindible aclararle esos objetivos a cada nuevo informante que se conoce. Si éste no es un informante potencialmente fundamental, puede bastar con transmitirle una breve reseña de los propósitos del estudio, pero en general es conveniente brindarle una explicación completa de los objetivos de la investigación a cada nuevo informante que se conoce, porque no se puede prever con exactitud, al principio, qué informantes llegarán a ser fundamentales, más adelante en el curso del estudio. Es útil que el investigador se ayude casi de memoria que se llevarán a cabo y de las medidas adoptadas para garantizar la máxima reserva y el mínimo riesgo. Si no es posible asegurar que la información será tratada como confidencial, el informante debe saberlo. Si el investigador va a estar presente en una escena dada, en el rol de observador participante, los informantes deben saber que en esa escena sus acciones y comentarios estarán bajo registro, aunque el investigador no tome notas ni realice grabaciones. Si los informantes desean que sus acciones queden fuera de registro, deben comprender claramente que sus acciones serán registradas al observarlos participante.

La preocupación de los informantes por el carácter evaluador de la investigación específica del observador es muy lógica, dada la ubicuidad de las observaciones con fines de evaluación en las escuelas. En última instancia, los propósitos del investigador también son, sin duda, evaluadores. Para describir las acciones de los individuos en informes narrativos

El investigador puede prever que los informantes someterán a prueba sus declaraciones acerca de la garantía de que mantendrá la reserva, el carácter acético de su perspectiva y otras consideraciones éticas que hayan sido negociadas. Esta puesta a prueba generalmente tiene lugar en las etapas iniciales de la relación con un nuevo informante. El informante puede pedirle un comentario crítico, o revelarle cierta información inofensiva y luego verificar si se ha corrido el rumor en la organización, para comprobar si el investigador le transmitió esa información a alguna otra persona del lugar.

Cuando se realiza una investigación en equipo es muy importante que todos los miembros del equipo cumplan estrictamente una regla básica: no hacer nunca comentarios a otros miembros del equipo acerca de ninguna cosa observada en el contexto mientras están en él. Las conversaciones privadas entre miembros del equipo de investigación en el lugar pueden ser escuchadas por los participantes, a veces con consecuencias desastrosas para la credibilidad del equipo.

Un excelente modo de crear y mantener un clima de confianza en el lugar es hacer intervenir directamente a los informantes en la investigación, en carácter de colaboradores del investigador. Al principio, los intentos del investigador de formar ese tipo de asociación con los informantes fundamentales pueden suscitar cierta desconfianza. Es posible que los informantes los interpreten como intentos de manipulación por parte del investigador, dado que la autopercepción de los profesores en especial, es que ellos no son expertos; si alguien es el experto en la asociación entre investigadores y docentes, el profesor puede presuponer que ese experto es el investigador. Esta percepción de los enseñantes puede persistir pese a los esfuerzos de los investigadores por refutarla. Pero, con el correr del tiempo, es posible desarrollar una genuina asociación en la que el profesor y el investigador comienzan a elaborar preguntas y a recoger datos en forma conjunta. (Sobre el proceso de la investigación de campo colaborativa en el terreno de la enseñanza, véase FROST y WALSH, 1980.)

Para desarrollar un entendimiento inicial, así como para forjar una relación de cooperación con los informantes fundamentales, el investigador debe tener una idea clara de las principales preguntas que guiarán la investigación y de los procedimientos de recopilación de datos que podrán emplearse para seguir las directrices marcadas por las preguntas orientadoras del estudio. Esto presupone la adhesión a la concepción deliberativa del proceso de la investigación de campo, antes mencionada en la presente exposición.

La recopilación de datos como proceso de indagación

La indagación comienza en el campo con las preguntas de investigación que guían el estudio. Hay tres aspectos cruciales en el inicio: identificar toda la gama de variación en los modos de organización social formal e informal (relaciones de rol) y las perspectivas de significado; b) registrar sucesos reiterados de acontecimientos a través de una amplia gama de acontecimientos en el lugar, para que más tarde pueda establecer el carácter típico o atípico de ciertos tipos de acontecimientos con su correspondiente organización social característica, y observar los acontecimientos que ocurren en cualquier nivel dado de sistema. (el aula, la escuela, el grupo de lectura) en el contexto de los acontecimientos que ocurren en los niveles inmediatamente superior e inmediatamente inferior. Esto significa que si se están observando los grupos de lectura, porque un problema de la investigación llevó al investigador a centrarse en los aspectos relativos a la enseñanza de la lectura, se observarán al menos los acontecimientos que conforman el acontecimiento que es el grupo de lectura (series de turnos de intervención relacionados temáticamente, movimientos del profesor, movimientos de los alumnos) y también acontecimientos en el nivel del aula, como por ejemplo, las jerarquías de las posiciones formales e informales de los alumnos en el aula como totalidad, o los modos característicos que tiene el profesor de organizar la interacción con los alumnos en otros acontecimientos aparte de la lectura. Lo ideal es observar una gama aún más amplia de niveles del sistema, para detectar posibles conexiones de influencia; por ejemplo, influencias en el propio establecimiento, ejercidas por otros docentes y por el director, que podrían afectar la práctica docente del profesor e influencias procedentes de la vida de los niños fuera de la escuela, en sus hogares y en otros lugares de la comunidad, que podrían afectar sus acciones en el grupo de lectura. En el trabajo de campo nunca se considera un único nivel del sistema, aislado de otros niveles; este es un rasgo básico de la técnica sociocultural de la que derivaron los métodos observacionales participativos.

A fin de determinar toda la gama de variación en las disposiciones de la organización social, las perspectivas de significado y las condiciones de influencia dentro y a través de los niveles del sistema en el lugar y sus ambientes circundantes, es necesario comenzar la observación y las entrevistas del modo más amplio que sea posible. Una vez que se ha avanzado en el proceso de investigación, la observación se va centrando, en etapas sucesivas, en un número cada vez más restringido de aspectos.

La progresiva resolución de problemas propia del trabajo de campo determina un proceso de muestreo secuencial. Dada la perspectiva amplia que se adopta al principio se puede creer que este proceso consiste en observar sin ninguna concepción previa, pero ésta sería una carnificación engañosa. Las concepciones previas y las preguntas orientadoras están presentes desde el principio, pero en ese momento el inves-

Hecho: un problema sobre exactamente adónde podrían llevarlo las presiones sociales. En consecuencia, empieza por realizar el muestreo más amplio posible del contexto y sus ambientes circundantes. Concretamente, esto significa que el investigador proyecta, en forma deliberada, un cierto tiempo en determinados lugares, en determinados momentos. Por ejemplo, al estudiar un aula, se comienza por buscar un sentido global de la comunidad vecinal de la escuela, recogiendo información por escrito sobre la comunidad escolar (por ejemplo, datos de censos), caminando y recorriendo en coche la comunidad, y entrando en las tiendas locales. Luego se accede al aula, a la que se observa durante todo el día, desde antes de que lleguen los alumnos hasta el momento de retirarse éstos. Una vez que ha identificado toda la gama de acontecimientos ocurridos durante el día y que, a través de repitidas observaciones, ha comenzado a determinar la frecuencia relativa con que se producen los diversos tipos de acontecimientos, el investigador puede empezar a formular hipótesis de aquellos que son de principal interés para el estudio. Comenzará entonces a restringir la gama de momentos y lugares de observación. Periódicamente, sin embargo, es posible que deba volver a un muestreo más comprensivo, a fin de recobrar la amplitud de perspectiva y de recopilar más ejemplos de casos a través de todo el espectro de acontecimientos sucedidos en el contexto. Esto le suministra un resplando adicional para las afirmaciones que luego quiera efectuar acerca del carácter típico y atípico (elevada y escasa frecuencia) de ciertos tipos de acontecimientos, o de ciertas reacciones de rol dentro de un acontecimiento o de una gama de acontecimientos.

A medida que el investigador se va centrando en un gama más restringida de acontecimientos dentro del contexto, también comienza a tener posibles conexiones de influencia entre éste y los ambientes que lo rodean. A diferencia del procedimiento propio de la etnografía comunitaria estándar, por el que se comienza con toda la comunidad como campo de análisis y se pasa, progresivamente, a investigar subunidades dentro de la comunidad, en el trabajo de campo educacional general, no se pasa con relativa rapidez, tras un ligero examen general de la comunidad al estudio continuo y concentrado de un entorno educativo específico por ejemplo, el aula o la clase de matemáticas). Después de haber trabajado el escenario principal con bastante profundidad, el investigador pasa otra vez a investigar los medios circundantes. La tarea analítica consiste en rastrear líneas de influencia fuera del aula y en los ambientes que la circundan. La clave para detectar estas líneas de influencia es que prescriben determinadas acciones dentro del aula y en los alrededores de los miembros del lugar educativos, alumnos, docentes y profesores de los miembros del lugar educativos, alumnos, docentes y profesores de los miembros de su vida fuera del contexto inmediato que influyen en la que se vive en él. Por lo general, los informantes no son totalmente conscientes de la simplicidad y la profundidad de estas influencias, en las que se incluyen supuestos culturalmente aprendidos e internalizados acerca de la naturaleza del currículo de mantener las relaciones sociales, del contenido de

las materias escolares, de la naturaleza humana y de las actitudes que determinan la delimitación de cada uno sobre la forma que podrían asumir aspectos tales como el trabajo, el juego, la fiabilidad, la capacidad económica y demás, al manifestarse en la vida cotidiana del aula.

Con el correr del tiempo, el investigador de campo va teniendo ideas que cada vez más claras respecto de los fenómenos más pertinentes para el estudio. En las etapas finales de la investigación, el foco de su atención puede ser muy restringido, en tanto las preguntas de investigación y las hipótesis de trabajo se hacen cada vez más específicas. El proceso de la investigación de campo como indagación deliberada ha sido descrito por algunos antropólogos (como AGAR, 1980; DOUBERT, 1982; DORF-BREWITE, 1984; GOETZ y LECAHIRE, 1984; LEVINE y OTROS, 1980; PÉLITO y PÉLITO, 1977), y por numerosos sociólogos (como GLASER y STRAUSS, 1979; LULI y VO, 1976; SCHLITZMAN y STRAUSS, 1974). Para los estudios sobre la enseñanza, este proceso ha sido descrito por ERICKSON (1973, 1977) y por MURRAY (1979), entre otros. Por razones de límites de espacio, no se tratará aquí una exposición completa del tema, pero nos parece oportuno tratar en mayor profundidad la índole del proceso de indagación en sí mismo.

El proceso racional limitado, de la resolución de problemas en el trabajo de campo

En el trabajo de campo, el investigador procura llegar a comprender acontecimientos cuya estructura es demasiado compleja para ser captada rápidamente, dados los límites de la capacidad humana para procesar información. Estos límites —lo que SMOLIN (1957) denomina *racionalidad limitada*— se compensan, en la observación participativa, a través de la presencia prolongada en el contexto.

El observador presente en determinados espacios y momentos en el lugar espera a que se sigan produciendo ciertos tipos de acontecimientos consistentes (disputas sobre la tenencia de tierras, decesos, nacimientos, preparación de la principal comida del día, atención del próximo enfermo en la oficina de desamparo, asistencia a una clase de lectura, interacción de los turnos para leer en voz alta, entrega de un ejercicio por escrito). El investigador puede buscar los lugares específicos dentro del contexto en los que hay más probabilidades de que se produzcan un determinado tipo de acontecimiento. Esto puede al observador en una situación análoga a la del sujeto de un experimento de aprendizaje, pues le brinda la oportunidad de hacer distintos ensayos para la ejecución de una tarea reiteradamente presentada. En este caso, las tareas es la de aprender cómo observar analíticamente un determinado tipo de acontecimiento y cómo llevar registros (notas de campo, grabaciones) de las acciones que ocurren en los acontecimientos, con el fin de estudiarlas con mayor detenimiento más adelante.

En cada ensayo de la observación de un acontecimiento reiterado,

el observador participante puede alterar levemente el foco de atención analizando atendiendo cada vez algunos aspectos de lo que sucede y desatendiendo otros. El observador también puede variar el foco de atención cuando, al leer las notas tomadas durante el acontecimiento y cuando las pasa en limpio después de haber concluido las observaciones del día. Un principio fundamental es que esta tarea de reflexionar y redactar las notas que suele llevar por lo menos la misma cantidad de tiempo que el dedicar a la observación, debe completarse antes de volver al lugar para continuar haciendo observaciones. Esto implica que el investigador debe prever que va a pasar cierto tiempo escribiendo notas; un día completo de observación en el aula debería ir seguido de un período de un día (o una noche) entero dedicado a la escritura. La redacción de las notas estimula la memoria y le permite al investigador agregar información a la contenida en las anotaciones iniciales. También estimula la integración analítica y la reflexión sobre los aspectos teóricos y la viabilidad de los pertinentes. No hay nada, durante el trabajo de campo, que pueda reemplazar a la reflexión que posibilita el tiempo dedicado a redactar las anotaciones originalmente tomadas, a redactarlas de manera más completa, incorporándoles percepciones analíticas.

Pese a los límites de la capacidad de procesar información, la observación y la reflexión prolongadas le permiten al observador desarrollar un modelo interpretativo de la organización de los acontecimientos observados. Tales modelos se construyen progresivamente a través de una serie de observaciones parciales, en un proceso análogo a un experimento de aprendizaje en el cual al alumno se le presenta una serie de ensayos. Hay dos grupos de decisiones del investigador de campo, acerca de los procedimientos a emplear que tienen especial importancia con el fin de corregir lo que tradicionalmente se denomina un sesgo en el muestreo y en la observación: a) las decisiones que toma el observador acerca de donde estar, en el espacio y en el tiempo, en el contexto, y b) las decisiones que toma el observador acerca de los focos de atención en cada ocasión. Las primeras decisiones conciernen al examen global de acontecimientos que realiza el observador participante. Las segundas atañen a la profundidad y la adecuación analítica de las observaciones acumulativamente efectuadas en una serie de ensayos.

Una de las principales ventajas de la observación participante es la oportunidad de aprender a través de la participación activa; el investigador puede poner a prueba su teoría de la organización de un acontecimiento ensayando diversos tipos de participación en el mismo. Una limitación importante en el trabajo de campo es la parcialidad de la visión de cada acontecimiento singular. Existe, en consecuencia, una tendencia a sesgar el muestreo que favorece a los tipos de acontecimientos que ocurren con mayor frecuencia, ya que ellos son los que el investigador busca a comprender más cabalmente con el correr del tiempo.

Analizaremos las consecuencias de la parcialidad en la próxima sección de este artículo. La desviación hacia lo típico está presente en la observación de campo también en otro sentido. Dadas las limitaciones,

sobre lo que se puede contemplar con mayor atención en cada ensayo observacional, es posible que el observador tienda a centrarse, en sí desde el principio, en alguna teoría de la organización que está incluyendo. En tal caso, durante sus observaciones le prestará atención sobre todo a aquellos aspectos de la acción que *confirman* la teoría inducida, pasando por alto otros aspectos que podrían anotarse como datos según los cuales dicha teoría quedaría refutada. Así, las pruebas que puedan inducir a refutarla tienen menos probabilidades de ser registradas en las notas de campo que las que puedan llevar a confirmarla.

La tendencia del investigador a apresurarse a sacar conclusiones inductivamente, desde las etapas iniciales del proceso de investigación, puede denominarse el *problema de la tipificación prematura*. Este problema hace que sea necesario realizar búsquedas deliberadas para detectar pruebas de refutación bajo la forma de casos discrepantes, o sea, ejemplos del fenómeno de interés cuya organización no concuerda con los términos de la teoría inducida. (Sobre la importancia del análisis de casos discrepantes, véase MEHAN, 1979, y la clásica exposición de LYONSVILLE, 1974.)

Otra manera de reducir las posibilidades de incurrir en tipificaciones prematuras o en desviaciones hacia el análisis de los acontecimientos constantes a expensas del de los esporádicos, consiste en hacer grabaciones durante el proceso de la investigación. Las grabaciones magnéticas o audiovisuales de acontecimientos recurrentes e infrecuentes en el escenario y sus ambientes circundantes le brindan al investigador la oportunidad de volver a presenciar, en fecha posterior, los acontecimientos grabados. Las grabaciones de la interacción que ocurre naturalmente en los acontecimientos no son sustituidas de la observación participante directa y el registro mediante anotaciones de campo. Pero esas grabaciones, sujetas a un análisis sistemático, pueden constituir una valiosa fuente adicional de datos en la investigación de campo. Con respecto a comentar brevemente la particular índole de las grabaciones y su análisis en parte, al que se efectuará en la próxima sección al análisis de datos, pero es conveniente destacar aquí ciertas diferencias con respecto a los tipos de resolución progresiva de problemas que posibilita la observación participante directa en el campo. La utilización de grabaciones como principal fuente de datos en la investigación de campo ha sido denominada «microetnografía» por ERICSSON (1975, 1976, 1982 a), «etnografía constitutiva» por MERTON (1979), y «microanálisis sociolingüístico» por GUYMERZ (1982). El proceso de la investigación microetnográfica fue descrito en detalle por ERICSSON y SHULTZ (1977/1981), por ERICSSON WILSON (1982), y por ERICSSON (1982 a).

Las grabaciones y su análisis difieren de la observación participante en un aspecto crucial. A diferencia del observador participante, el analista de registros documentales audiovisuales o magnetofónicos no espera en el contexto a que se produzca la ocurrencia de un determinado tipo de acontecimiento. Al revisar las pruebas documentales grabadas,

el analista está libre de las limitaciones que le impone al observador participante su inserción en el evento sociocultural de acontecimientos, en el tiempo y el espacio reales. El investigador hace un listado de todo lo grabado, identificando los principales acontecimientos por su nombre (por ejemplo, clases, reuniones, recreos, encuentros entre padres y docentes) e identificando también la presencia de informantes secundarios en ciertos acontecimientos. Luego revisa de principio a fin todo el conjunto de grabaciones a fin de encontrar ejemplos de acontecimientos frecuentes e infrecuentes, describiéndolos, por así decirlo, a través del tiempo y el espacio para identificar casos análogos. Esto es similar al examen inicial del lugar y su ambiente circundante que efectúa el observador participante, cuando debe decidir dónde situarse, en el tiempo y en el espacio, derivado del contexto. A continuación, el investigador identifica una determinada serie de casos, a partir del conjunto grabado, que son de especial interés. A esta altura, el investigador puede volver a presenciar esta serie de casos, pasando las correspondientes grabaciones. La posibilidad de volver a presenciar el mismo acontecimiento con el fin de realizar reiteradas observaciones es la principal innovación posible, tanto por el empleo de grabaciones en la investigación de campo. Esta innovación presenta claras ventajas y también limitaciones.

La primera ventaja es la posibilidad de efectuar un análisis completo. Debido a la oportunidad (técnicamente) limitada de volver a presenciar un acontecimiento, pasando la grabación correspondiente, ese acontecimiento puede observarse desde diversos focos de atención y perspectivas de análisis. Esto permite una descripción mucho más exhaustiva que las que puede elaborar el observador participante a partir de sus intuiciones de campo.

La segunda ventaja es la posibilidad de reducir la dependencia del observador de la tipificación analítica rutinaria. Dado que cada acontecimiento se puede volver a ver a través de la grabación, el observador tiene la oportunidad de deliberar. Puede dejar en suspenso la formulación de juicios interpretativos de las funciones (significados) de las acciones observadas. En la observación participante, muchas veces se realizan deducciones interpretativas que pueden ser falsas, un especial problema las etapas iniciales de trabajo de campo.

En el análisis microetnográfico de una película o una grabación de video, la oportunidad de mirar y escuchar más de una vez disminuye la función analítica. Esta independencia de los límites que le impone el tiempo real a la observación produce una profunda diferencia cualitativa entre el discurso oral y el escrito, en tanto el segundo se presta a ser revisado y también planificado en formas mucho más detalladas que el primero. Sobre esta diferencia, véanse Cooley, 1977 y Oss, 1977.

Con respecto a la observación de dependientes de datos, el análisis de grabaciones es el mejor de las fuentes de datos. Al analizar una grabación, en especial si es audiovisual, se pueden estudiar de manera bastante completa los acontecimientos poco frecuentes pasando varias veces la grabación correspondiente. Esta posibilidad no está al alcance del observador participante al tratar acontecimientos infrecuentes.

Existen dos limitaciones principales en el uso de grabaciones como principal fuente de datos. La limitación fundamental es que al pasar una grabación, el analista sólo puede intervenir de forma transitoria en el acontecimiento grabado. Así, la posibilidad de la experiencia vicaria de también su mayor limitación. El investigador no tiene la oportunidad de verificar las teorías que va elaborando mediante el recurso de poder observar en tanto participante activo en las escenas observadas. Esta oportunidad es uno de los rasgos distintivos de la observación participante.

Otra limitación radica en que, para encontrarle el sentido al material grabado, el analista normalmente necesita tener acceso a cierta información contextual que no brinda la propia grabación. El acontecimiento grabado se inserta en una diversidad de contextos: en las historias de vida y los vínculos sociales de los participantes, y en las circunstancias sociales más amplias que afectan a los acontecimientos, incluyendo la importancia que tiene la pertenencia de los participantes a un grupo étnico, cultural o de clase social respecto de los participantes a un grupo su conducta conjunta en el acontecimiento grabado. El hecho de conocer los modos en que el acontecimiento que ocurre cara a cara se inscribe en la red de influencias existente dentro del contexto y entre el contexto y sus ambientes más amplios (incluyendo la sociedad total en la que éste está situado) puede ser decisivamente importante para ayudar al investigador a lograr una comprensión interpretativa de la organización inmediatamente local de la interacción incluida en el acontecimiento, o una comprensión interpretativa de la significación que encierra el hecho de que un acontecimiento haya ocurrido de un modo determinado y no de otro modo posible. Algunas veces, los contextos más amplios que rodean a un acontecimiento pueden no tener mucha influencia sobre el desarrollo de las relaciones sociales cara a cara dentro del acontecimiento. En esos casos, el análisis no se ve invalidado por la ausencia de un marco etnográfico más amplio. Pero cuando el mundo circundante tiene una fuerte influencia sobre las escenas inmediatas de la interacción cara a cara en el aula, y cuando esa influencia determina una elección sistemática sobre el desarrollo de las relaciones en el aula, entonces la ausencia de un marco contextual más amplio procedente del trabajo de campo etnográfico general puede invalidar el análisis.

Otras limitaciones de la microetnografía —la ausencia de la participación como medio de aprendizaje y la ausencia de información contextual más allá del marco de la grabación— pueden ser superadas mediante la combinación de la etnografía regular y la microetnografía.

III SEGUNDA PARTE

Esta segunda parte de la antología incluye cuatro textos que se han seleccionado por que abordan distintos instrumentos de recolección de información; el primero de ellos es un artículo elaborado por el Consejo estatal Técnico de la Educación en Veracruz, bosqueja una concepción de evaluación y a partir de ella ofrece una serie de fichas que aportan elementos básicos para la evaluación de distintos aspectos del trabajo del profesor como son la evaluación diagnóstica, el seguimiento del desarrollo de habilidades, la evaluación formativa mensual, la evaluación sumativa, la autoevaluación, etc.

El texto de Karmel ofrece una serie de consideraciones sobre los instrumentos que elaboran los propios profesores para evaluar el aprendizaje de sus alumnos; el autor cuestiona la capacidad de los profesores para elaboración de este tipo de instrumentos, reconociendo, al mismo tiempo, que es necesario que los hagan, por ello ofrece recomendaciones para que los instrumentos que sean diseñados cubran los requerimientos mínimos generales. Es posible que no compartamos la opinion del autor sobre los docentes, sin embargo sus señalamientos resultan provechosos para la elaboración y aplicación de estos instrumentos.

El texto de Galindo Chavanel es una ficha sobre un instrumento para registrar los datos dentro de un proceso de observación, el modelo que ofrece puede ser facilmente adaptado a diferentes propósitos. La autoria maneja una serie de pautas con sus respectivos indicadores, siguiendo el esquema presentado usted puede elaborar sus propias categorías observacionales; revise con cuidado las pautas que presenta para observar el aspecto emocional ya que la mayoría de ellas requieren de mayor precisión, la autora maneja terminos muy generales que pueden ser interpretados de distintas formas por distintos observadores, tome en cuenta esta situación y, en el caso que su trabajo lo requiera, trate de que los criterios de observación queden claros y sea fácil su interpretación.

Por último se presenta un texto que es bastante extenso, se decidió incluirlo por que presenta de manera amplia y con bastantes varios instrumentos para la recuperación de información, seguramente usted no utilizará todos ellos sin embargo es posible que las observaciones que ofrece le serán de utilidad; un punto que resulta conveniente revisar, independientemente del tipo de instrumento o estrategia que haya seleccionado, es el referido a el desarrollo de ítems. Identifique cuales son los instrumentos que incluye el texto y tengalos presente por si necesita recurrir a ellos.

INTRODUCCIÓN

El componente 11 del Programa para Abatir el Rezago Educativo (PARE) incluye, entre sus líneas de acción, la correspondiente a la Evaluación en el Aula. El objetivo general de ésta se refiere a la implantación de estrategias de apoyo a la docencia para mejorar los procesos de evaluación en su principal ámbito de trabajo: el aula.

La metodología empleada para el diseño de estrategias a implantar, contempló el diseño y aplicación de un cuestionario a una muestra representativa de docentes en las cuatro entidades consideradas por el PARE (Chiapas, Guerrero, Hidalgo y Oaxaca). Los contenidos del cuestionario permitieron obtener información sobre las características de los procesos de evaluación que los docentes instrumentan en el aula. Esta información se procesó por medios computarizados y fue posteriormente analizada para obtener un diagnóstico que, a su vez, fundamentara el diseño de estrategias viables, válidas y útiles para apoyar la labor docente en lo que se refiere a la evaluación en el aula.

Este documento presenta las estrategias de apoyo emanadas del diagnóstico, constituidas por sugerencias prácticas que sin duda se verán enriquecidas con la experiencia que, según se constató mediante el cuestionario aplicado, tiene la gran mayoría de los docentes de Educación Primaria, en las entidades participantes.

Es importante mencionar que independientemente de esta rica experiencia, los docentes en general manifestaron carencias en cuanto a la actualización en materia evaluativa y se observó una marcada tendencia hacia la práctica tradicional de la evaluación donde es privilegio la asignación de calificaciones.

Asimismo, la información obtenida indica en lo general una acusada utilización del sentido común en la evaluación dentro del aula, así como una falta generalizada de sistematización de este proceso.

Las estrategias diseñadas pretenden apoyar la

actividad del docente en el aula mediante la sugerencia de formas concretas para sistematizar y explotar, en lo posible, la evaluación en el aula, de tal modo que constituya un mecanismo de retroalimentación del proceso de Enseñanza- Aprendizaje.

Las estrategias propuestas responden a la teoría que considera que el niño construye su propio conocimiento con base en sus experiencias con el entorno, esta teoría se llama constructivista de desarrollo cognoscitivo.

El presente texto está integrado por un breve apartado inicial que tiene los principios conceptuales que enmarcan las estrategias sugeridas más adelante. Su brevedad responde a la exigencia de operatividad del documento, cuya intención es facilitar el trabajo del profesor. Se incluyen, además, doce fichas técnicas en cada una de las cuales se desarrolla una estrategia de evaluación. Cada ficha contiene cuatro apartados: la situación o problema de evaluación, la estrategia propuesta, el apoyo teórico indispensable y la bibliografía mínima.

Las fichas proponen estrategias alternativas que tocan los temas sobre la planeación y acreditación de la evaluación; la autoevaluación y evaluación grupal; el seguimiento del desarrollo de habilidades y la investigación en el aula como formas de evaluar; las evaluaciones diagnóstica, formativa y sumativa; la evaluación de los contenidos por medio de exámenes a libro abierto, grupales y globales para resolver en casa; y la observación y las listas de cotejo como la técnica y el instrumento idóneos para evaluar. Se incluyen, al final, la bibliografía general y los anexos.

Es importante señalar que esta gama de estrategias representa únicamente una alternativa más para el profesor. Es él, y sólo él, quien habrá de seleccionar aquella o aquellas (se mencionen o no en este documento) que respondan a sus necesidades y características docentes, así como a las de su grupo. En ningún momento se pretende que el profesor las aplique todas.

SOBRE ALGUNOS CONCEPTOS

La concepción de que el niño construye su propio conocimiento requiere que los profesores tengan una actitud distinta ante la conducción del proceso Enseñanza-Aprendizaje; asimismo, la disposición y acciones necesarias para adquirir una formación teórica que propicie y apoye dicha actitud.

La actitud del profesor deberá ser congruente con la concepción de aprendizaje que otorga gran importancia a la acción de los niños sobre el objeto del conocimiento (experiencia con su entorno). Dicha acción debe ser física al interactuar con los objetos de estudio (contenidos), y mental al reflexionar sobre dichos objetos sobre lo que se hizo con ellos.

Así:

UN NIÑO QUE
APRENDE
A DIVIDIR

OBSERVA EL PROCESO DIVIDIR
• CON OBJETOS
• CON NÚMEROS
DIVIDE OBJETOS

DIVIDE NÚMEROS
• NO PUEDE DIVIDIR
• TRATA DE RECORDAR
EL PROCESO QUE OBSERVÓ
• RECONSTRUYE LAS ACCIONES Y
LAS REFLEXIONES QUE HIZO

INTERACTÚA
CON OBJETO
DIVISIÓN

REFLEXIONA
SOBRE ÉSTE

CONSTRUYE
SU PROPIO
ESQUEMA
SOBRE LA
DIVISIÓN

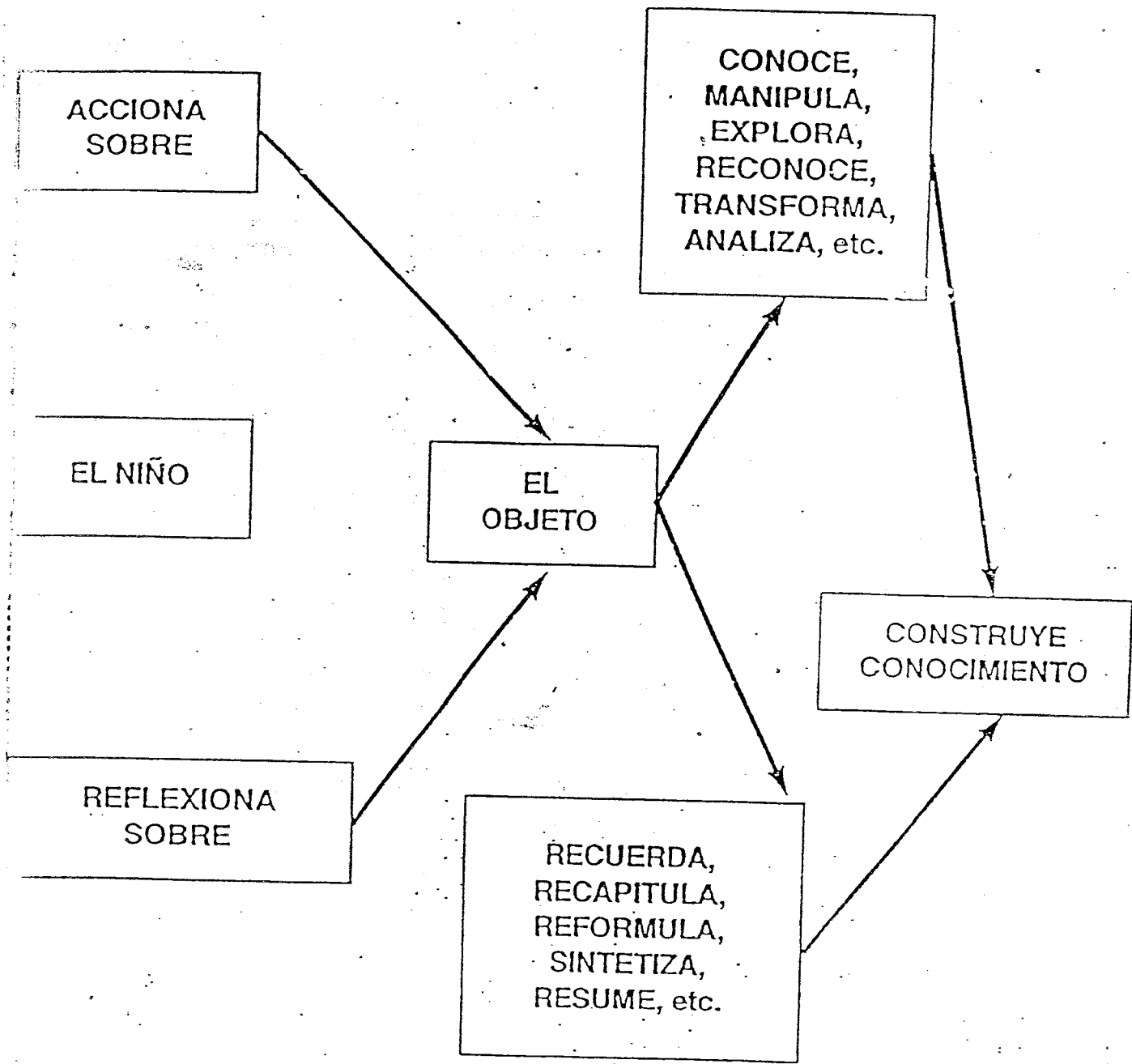
Como se puede observar, el objeto es siempre la parte de la realidad que el niño está aprehendiendo, haciéndola suya, apropiándose de ella. Es la parte de la realidad que él está contruyendo a partir de su propia experiencia.

EL PROFESOR
AL ENSEÑAR
A DIVIDIR:

EL PROFESOR
ESTRUCTURA EL MEDIO AMBIENTE
• REÚNE MATERIALES
• EXPLICA CÓMO DIVIDIR NÚMEROS
• PRESENTA EJERCICIOS Y AYUDA
A RESOLVERLOS
• OBSERVA LOS PROCESOS DE
APRENDIZAJE EN LOS NIÑOS
• REGISTRA SUS AVANCES, LOGROS
Y OBSTÁCULOS
• REESTRUCTURA EL MEDIO AMBIENTE

EL NIÑO
CONSTRUYE
SU PROPIO
ESQUEMA

Por tanto son dos los principios que deben normar la construcción de todo aprendizaje que se pretenda sea significativo:



Concebir así el aprendizaje implica cambiar nociones y conceptos que hasta el momento se habían manejado. Dicho cambio implica reorientar la práctica docente. Es decir, que sea el profesor mismo quien reconozca cuáles de sus actividades requieren cambios, ya sea porque no funcionan ante nuevos retos (avance del conocimiento, problemas sociales, etc.) o porque el profesor y/o la comunidad escolar consideren viable y pertinente conducir el proceso de aprendizaje con otro enfoque. Para una explicación más amplia se exponen en continuación las funciones del docente dentro de la concepción teórica que se plantea.

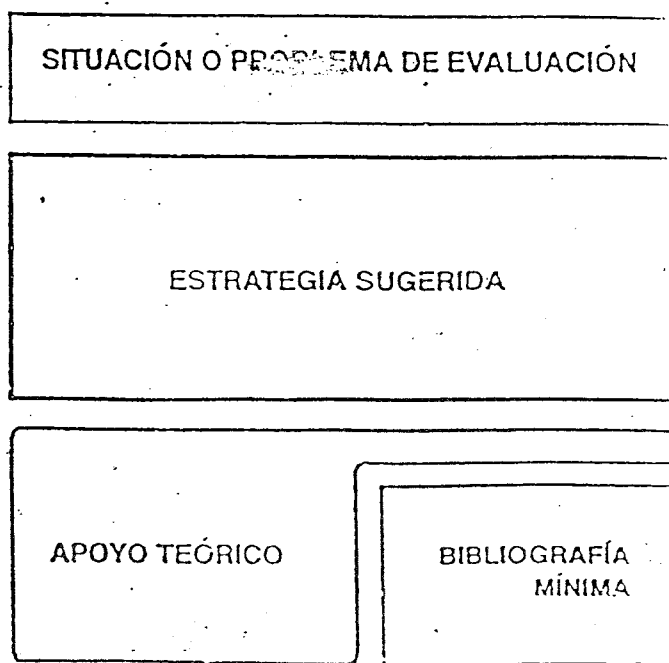
FUNCIONES DEL DOCENTE

- Planea el contenido de sus clases para propiciar la adquisición de conocimientos y el desarrollo de habilidades.
- Conoce a fondo la materia de estudio.
- Tiene un panorama claro del desarrollo psicológico del niño y es capaz de identificar estadios por la observación y su registro.
- Es guía, asesor y planeador de procesos de aprendizaje. Los contenidos son un medio.
- Conoce las bases teóricas y es experto en la selección y utilización de material extraído del medio circulante para propiciar situaciones graduadas que generen aprendizajes significativos.
- Habla cada vez menos y hace intervenir cada vez más a los alumnos.
- Observa, registra y estudia cuidadosamente las situaciones educativas para su adaptación inmediata o por su relación con futuros procesos de desarrollo.
- Evalúa procesos del desarrollo de habilidades que se plantea en el enfoque de los programas.
- Propicia y estimula la autoevaluación y evaluación en equipo y las considera para la acreditación.

Como ya se había señalado, desempeñar cada una de estas nuevas funciones requerirá que el profesor construya una formación teórica para el adecuado desempeño de cada una. El formarse teóricamente, si bien requerirá de mucho tiempo, puede ser fortalecido por la experiencia, la cual es una base importante y constituye, de hecho, el cimiento de una formación que debe ser paulatina y llevarse a cabo en la propia práctica docente.

Uno de los campos donde el profesor habrá de formarse es, sin duda alguna, el de la evaluación del aprendizaje, dada su importancia como elemento retroalimentador tanto del aprendizaje como de la planeación docente.

A continuación se presentan algunas sugerencias para desarrollar el trabajo de evaluación dentro del aula. Es importante considerar que el análisis de cada profesor realice de los programas del grado a atender tiene una singular importancia para determinar tanto los conocimientos a impartir como las habilidades a desarrollar y, por tanto, será un instrumento esencial para poner en práctica algunas de las estrategias que se proponen; el Anexo 1 presenta un ejemplo del análisis de programas que requiere realizar). Por otra parte además de formular estas propuestas se pretende sugerir estrategias que brinden un apoyo teórico a fin de coadyuvar a la formación teórica aludida. La forma de presentar estas estrategias trata de ser fluida y de fácil lectura y consulta. Cada estrategia se diseñó en una ficha, esta numerada y cada una tiene la siguiente estructura: se expone en primer término la situación o problema de evaluación; se desarrolla en seguida la estrategia sugerida para incluir después el sustento teórico y la bibliografía mínima sobre éste último. El siguiente gráfico esquematiza la estructura de las fichas:



LA EVALUACIÓN A PARTIR DE UN PLAN ANUAL (FICHA 1)

La evaluación se ha manejado tradicionalmente como una actividad a realizar periódicamente (por lo general al final de mes y al final de año): No obstante, es una actividad cotidiana que los maestros realizan intuitivamente. No deriva de un plan previamente elaborado, lo cual genera que se realice de manera azarosa. La evaluación debe ser una actividad constante, permanente, que permita, además de estimar los avances de los niños, tomar medidas correctivas en caso necesario, así como planificar nuevas situaciones de aprendizaje.

1. Lo más importante de un Plan de Evaluación es elaborar por escrito la concepción que el profesor tiene, desde un punto de vista teórico, de lo que es la Evaluación. Para ello, debe redactar un documento sencillo en el que establezca una idea de evaluación y sus finalidades, debe partir necesariamente de una definición de aprendizaje.

2. El profesor debe tener un plan de evaluación anual que atienda a tres aspectos: el aprendizaje de los alumnos, la actividad docente y el grupo. Esto le permitirá tener una visión de las evidencias que podrá reunir.

El plan debe describir cómo evaluará a los alumnos (Tipos de Evaluación que utilizará, Estrategias de Evaluación, Instrumentos de Evaluación, Momentos y Fechas); cómo evaluará su propia actividad (Autoevaluación, Evaluación Externa y Evaluación de Grupo) y cómo realizará la evaluación del grupo (Tipos, Rutinas, Instrumentos, Momentos y Fechas).

El Plan Anual de Evaluación (en sus tres aspectos: alumnos, maestros y grupo) debe ser un documento sencillo, orientador de la actividad del profesor y manejable. A continuación se sugieren tres formatos para dicho plan, elaborados a manera de cuadros.

PLAN DE EVALUACIÓN ANUAL DE LOS ALUMNOS

TIPOS DE EVAL.	PROCESOS A EVALUAR	ESTRATEGIAS DE EVALUAC.	MOMENTO	INSTRUMENTOS	FECHA
----------------	--------------------	-------------------------	---------	--------------	-------

PLAN DE EVALUACIÓN ANUAL DEL PROFESOR

ASPECTOS A EVALUAR	AUTOEVALUACIÓN	EVALUACIÓN DEL GRUPO	EVALUACIÓN EXTERNA
--------------------	----------------	----------------------	--------------------

PLAN DE EVALUACIÓN ANUAL DEL GRUPO

TIPOS DE EVAL.	PROCESOS A EVALUAR	RUTINAS	INSTRUMENTOS
----------------	--------------------	---------	--------------

Los resultados obtenidos al llevar a cabo el o los planes de evaluación permitirán reunir las evidencias suficientes para emitir una evaluación cualitativa. Por ejemplo: la información obtenida mediante el Plan de Evaluación de los Alumnos permitirá establecer en qué procesos se han desarrollado todos o individualmente y con qué profundidad.

Las evidencias de evaluación se irán reuniendo en los instrumentos destinados para ello.

...la evaluación del proceso enseñanza aprendizaje no se inicia cuando termina el curso, o cuando comienza a realizarse los exámenes. Antes de empezar el curso cuando se especifican sus propósitos, es cuando en realidad comienza el proceso de evaluación lo cual implica que la evaluación no es una etapa fija ni final del proceso docente. En la planeación de un curso... va ya implícita la consideración del proceso de evaluación. (MORÁN, 1981, p. 24)

MORÁN OVIEDO, FORTINHO. "La Evaluación de los Aprendizajes y sus Implicaciones Educativas y Sociales", en Perfiles Educativos No. 13. México, CISE-UNAM, 1981, p. 21-36

LA EVALUACIÓN EN EL AULA COMO INVESTIGACIÓN (FICHA 2)

La evaluación en el aula ha sido utilizada como un instrumento de medición de conocimientos y de acreditación escolar. Proporciona información que permite a los maestros percatarse de "cómo van los niños". Sin embargo, eso lo saben ellos sin necesidad de calificar porque la evaluación es una actividad cotidiana que llevan a cabo, si bien intuitiva y asistemáticamente, con juicios por lo general acertados.

1. El profesor identifica los conocimientos correspondientes al grado y las habilidades que se requiere desarrollar en el documento relativo al plan y programas de estudio de educación primaria (ver ejemplo en Anexo 1).
2. Se informa bibliográficamente sobre dichos procesos y redacta las características fundamentales de cada uno.
3. Elabora una evaluación diagnóstica sobre estos aspectos (ver ejemplo en la Ficha 4. Evaluación Diagnóstica Anual a Partir del Desarrollo de Habilidades) y la aplica a los alumnos al iniciar el año.
4. Con base en los resultados de la evaluación establece hipótesis de trabajo para su INVESTIGACIÓN ANUAL. Las hipótesis son supuestos sobre cómo considera (con base en su experiencia) que se van a desarrollar las habilidades requeridas, cuáles van a ser los obstáculos y problemas que encontrarán y cuáles las posibles soluciones. Por ejemplo:

La expresión oral de los niños de tercer grado de San Gregorio, Gro., es poco fluida. En general, utilizan tiempo pasado para expresar hechos en presente, empiezan a utilizar modismos de:

lugar sin saber su significado y en la mayoría de los casos no terminan las frases u oraciones (es una costumbre propia del lugar). El mayor problema que se enfrentará será lograr que expresen ideas completas cada que en sus hogares no se acostumbra. Para ello se propiciará que escriban descripciones de situaciones familiares y luego, como tarea, la lleven en un texto incompleto para ser completado por miembros de su familia. Para el manejo de modismos se propiciará que investiguen en su comunidad los diversos significados de cada uno para llegar a conocer el más usual. Éstas serán tareas cotidianas que se efectuarán, por lo menos una vez a la semana durante los primeros tres meses del año escolar.

5. Elabora una estrategia de investigación que incluya, por ejemplo:

HABILIDADES A DESARROLLAR	CARACTERÍSTICAS A OBSERVAR	TÉCNICAS DE OBSERVACIÓN Y REGISTRO	FECHAS
EXPRESIÓN ORAL	PLÁTICAS INFORMALES	LISTA DE COTEJO No. 1	Principio y Fin de año
	LECTURA ORAL	LISTA DE COTEJO No. 2	1 vez/mes
	DRAMATIZACIONES	LISTA DE COTEJO No. 3	1 vez/mes
	ANÉCDOTAS Y NARRACIONES	LISTA DE COTEJO No. 4	1 vez/mes

6. Por cada registro de observación que efectúa, el maestro redacta una apreciación de cómo se va desarrollando el proceso en los niños, qué problemas y qué logros se presentan. Asimismo, en qué medida se van comprobando o rechazando las hipótesis planteadas.

7. Al finalizar el año (o el tiempo que haya establecido para la investigación) redacta un reporte en el que debe plantear los resultados, las conclusiones y las recomendaciones para futuros casos.

"La investigación de la práctica educativa por parte del propio docente, es un proceso encaminado a la descripción, explicación y valoración de aquello que sucede en el aula... El profesor-investigador posee una condición de privilegio para representarse [elaborar un modelo de interpretación] la práctica educativa y es el mejor crítico interno del proceso enseñanza-aprendizaje." (GERSON, 1979. p. 4) "La obsesión por valorar la educación sólo a través de sus resultados inmediatos impide estudiar... aquellos efectos educativos que más profundamente arraigan en los sujetos.

Una pedagogía pragmática sólo se conforma con aquello que de manera inmediata puede corroborar. Pero estos resultados son básicamente aparentes y no dan cuenta de un conjunto de procesos que paulatinamente van estructurando al sujeto". (DÍAZ, 1987. p. 11)

DÍAZ BARRIGA, ÁNGEL "Problemas y Retos del Campo de la Evaluación Educativa", en *Perfiles Educativos* No. 37. México, CISE/UNAM, 1987. pp. 3-15

GERSON, BORIS. "Observación participante y diario de campo en el trabajo docente", en *Op. Cit.* 5. 1979. pp. 3-22

5. A primera vista se observa que Sergio y Pablo no han desarrollado esta capacidad y necesitarán un Plan de Trabajo para superarlo, los demás alumnos poseen algunos rasgos y otros no, de acuerdo con las características de su personalidad.

El profesor platica con Pablo y Sergio para formular un plan mensual que les ayude y permita aprender a participar, solicitar ayuda, compartir objetos y juegos, iniciar y/o dirigir juegos o pláticas con sus compañeros, etc.

6. El profesor registra en una libreta especialmente destinada para el "SEGUIMIENTO DE LOS ALUMNOS" tanto las observaciones como las medidas tomadas para desarrollar las capacidades de Pablo y Sergio. En esta libreta existe un espacio destinado para cada alumno. En ella se hacen anotaciones solamente cuando el alumno tiene algún obstáculo o problema en el desarrollo del aprendizaje. Ejemplo: Pablo (93.09.12) parece no tener desarrollada la capacidad para relacionarse con los demás.

Durante este mes él dirigirá el juego de la mañana ayudado por mí y revisará el aseo del salón; será el primero a quien se le revise la tarea y ayudará a sus compañeros a corregirla. Sergio (93.03.19) parece no tener desarrollada la capacidad para relacionarse. Efectuará las mismas actividades que Pablo y entre los dos me informarán cómo lo han hecho.

"Entendemos la evaluación como el estudio del proceso de aprendizaje en un curso, un taller, un seminario, etc; con el fin de caracterizar los aspectos más sobresalientes del mismo, y a la vez, los obstáculos que hay que enfrentar" (PANSZA, PÉREZ Y MORÁN, 1986, p.111). "La evaluación de procesos parte de estudiar las condiciones en que se desarrolla una acción educativa con la finalidad de imponer correctivos durante su ejecución". (DÍAZ, 1987, p.3)

PANSZA, PÉREZ Y MORÁN. Operatividad de la Didáctica. Tomo 2. México, Gemika, 1986, pp. 97-137

DÍAZ BARRIGA, ÁNGEL. "Problemas y Retos del Campo de la Evaluación Educativa", en Perfiles Educativos No. 37. México, CISEUNAM, 1987, pp. 3-15

EVALUACIÓN DIAGNÓSTICA ANUAL A PARTIR DEL DESARROLLO DE HABILIDADES

(FICHA 43)

La evaluación diagnóstica ha sido entendida como la detección de requisitos de conocimiento que se poseen para poder entender y trabajar ciertos contenidos. Sin embargo, en el caso particular de los niños, los requisitos no sólo son de conocimiento sino fundamentalmente de madurez. ¿Cómo diagnosticar dicha madurez?

1. Seleccionar los conocimientos correspondientes al grado y las habilidades que se requiere desarrollar.

2: Organizar un plan de actividades para una semana en el que se pueda observar el nivel de madurez que tienen los niños en las distintas habilidades. Se recomienda trabajar la primera semana del año en esta evaluación porque aportará elementos que permitirán planificar adecuadamente el nivel de profundidad que se va a tener en el trabajo de todo el año. Si bien parecería representar atraso de una semana en el programa, en realidad implica un avance notorio porque asegura una buena planificación.

El plan de actividades podría diseñarse así:

HABILIDADES A DESARROLLAR	LUNES
1. DIÁLOGO, EXPOSICIÓN DE ANÉCDOTAS Y EXPOSICIÓN DE TEMAS	Formarán equipos de trabajo para la semana o inventarán sus reglas.
2. EXPRESIÓN ESCRITA, ELABORACIÓN DE TEXTOS CON TEMA PRESCRITO	Escribirán individualmente un texto: "Mi opinión acerca de la Contaminación del Medio Ambiente".
3. DISFRUTE Y COMPRENSIÓN DE LA LECTURA	Leerán un texto sobre la importancia de la verdad y comentarán en equipo sus opiniones.
4. ARGUMENTAR Y EXPLICAR IDEAS SOBRE EL CUIDADO DEL MEDIO AMBIENTE	Comentarán en equipo algunas normas para el cuidado del medio ambiente.
5. FORMACIÓN DE ACERVOS	Consultarán la biblioteca escolar e investigarán las tradiciones mexicanas.

3. Diseñar un programa por día de tal forma que durante los 5 días hábiles de la semana se evalúen la totalidad de habilidades a desarrollar, consideradas por el docente después de analizar los programas del grado. En este caso se recomienda, también, observar con mayor detenimiento 10 niños y hacer el registro de las observaciones. Si bien se observa a los demás, no se hace registro de ellos.

Diseñar una Lista de Cotejo para registrar las observaciones que se hagan de los niños con los rasgos

que considera importantes (ver Anexo 2, Ejemplo de Lista de Cotejo y Ficha No.11, La lista de Cotejo como una forma de Registrar Observaciones a Evaluar).

Se sugiere que sobre cada rasgo sólo se registren tres o cuatro posibilidades o niveles, por ejemplo: D=detalladamente, S=someramente y N=NO LO HACE. Estas posibilidades o niveles pueden variar según el proceso y la actividad que se observe y dependerán del nivel de profundidad con que planificará las experiencias de aprendizaje durante el año.

Carlos Zarzar Charur señala que la evaluación es "el momento de recuperación de los aprendizajes; (y) al mismo tiempo... un momento más de aprendizaje... consiste en una reflexión sobre lo que se aprendió y sobre la manera en que se aprendió..." (ZARZAR, 1983. p. 45). Se podría inferir que la Evaluación cognoscitiva es también un momento para la recuperación de lo que ya se ha logrado y de lo que aún no, permitirá percatarse de qué habilidades se dominan y cuáles están en desarrollo y, por tanto, de lo que se deberá aprender en el futuro. Se busca que los alumnos no sólo expresen conocimientos que obtuvieron en su hogar o el grado inmediato anterior, sino que expresen cómo los manejan y cómo les han ayudado a avanzar en su desarrollo.

ZARZAR CHARUR, CARLOS. "Diseño de Estrategias para el Aprendizaje Grupal" en Perfiles Educativos No.1, Nueva Época. México, CISE/UNAM, 1983. pp.34-46 (abril-junio)

**EVALUACIÓN FORMATIVA MENSUAL
POR MEDIO DE ALUMNOS MONITORES (FICHA 5)**

La Evaluación Formativa es un problema para el profesor porque tiene grupos numerosos y no cuenta con formas de registro adecuadas.

1. Seleccionar conocimientos a aprender y habilidades a desarrollar por mes y por asignatura.
2. Elaborar una vez al mes un Cuadro de Relación y Secuencia entre desarrollo de habilidades, asignaturas y contenidos.
3. Elaborar una Lista de Cotejo con los rasgos a observar. Es importante que se haga una sola lista por meses, la cual deberá incluir como rasgos a observar aquellas expresiones del niño que evidencien o pongan manifiesto que ha trabajado un contenido o desarrollado una habilidad (ver Anexo 2, Ejemplo de Lista de Cotejo).
4. Observar a los alumnos en su desempeño a lo largo del trabajo durante el mes y seleccionar a los:

diez que lo hayan hecho mejor para que funcionen como monitores.

5. Destinar un tiempo en el último día de la tercera semana de cada mes para trabajar con ellos y enseñarles cómo se usan las listas de cotejo, asimismo qué significa cada uno de los rasgos.

6. Explicar que la semana próxima le ayudarán a evaluar el trabajo de sus compañeros registrando cómo éstos hacen algunos trabajos y ayudándoles a corregir errores o superar obstáculos.

SESIÓN DE EVALUACIÓN FORMATIVA DIARIA DURANTE LA CUARTA SEMANA DE CADA MES

El Maestro:

1. Diseña actividades globalizadoras de lo que se ha aprendido durante el mes, una por cada materia o grupo de materias.

2. Convoca y organiza la formación de equipos vigilando que cada equipo tenga un monitor.

Los Alumnos Monitores:

3. Se presentan con sus compañeros y explican el propósito de su trabajo.

4. Observan a sus compañeros efectuar la actividad y anotan sus observaciones en la Lista de Cotejo. Al finalizar la actividad, los monitores comentan los resultados con sus compañeros y resuelven dudas de ser posible.

5. Dirigen los comentarios finales sobre avances, logros, obstáculos y propuestas.

6. Diseñan con sus compañeros el Plan de Trabajo Individual Mensual para superar sus problemas (ver Anexo 3, ejemplo de Plan Individual de Recuperación Mensual).

7: Al terminar, los alumnos monitores se reúnen con el profesor para comentar sus resultados.

8. El profesor, con base en los resultados, reorienta su trabajo del mes siguiente.

NOTA: Se sugiere sea utilizado con alumnos de tercero a sexto grados. Cuide que los monitores no sean siempre los mismos alumnos, es necesario que todos aprendan a observar y evaluar, y entiendan la importancia de la evaluación para mejorar su aprendizaje.

La Evaluación Formativa a partir del proceso grupal se fundamenta en la idea de concebir al aprendizaje grupal "...como la estrategia metodológica idónea para desarrollar paralelamente y con sentido crítico y constructivo, los procesos de evaluación y acreditación... Se aprende a pensar en grupo; con otros se afrontan procesos de esclarecimiento tanto de aspectos relativos a una materia de estudio como de las dificultades y los problemas que implica este tipo de trabajo... es un proceso en el que los involucrados asumen alternativamente el rol de sujetos y objetos de evaluación ..." (PANSZA Y OTROS, 1986. pp.

100, 106 y 107). La Evaluación de procesos es una actividad que pueden realizar los alumnos a partir de la estimación de sus propios avances y los de sus compañeros. Les permite conocer qué saben y lo que no, cuáles son los obstáculos que encuentran y, por comparación con el trabajo de sus compañeros, saber cómo podrían superarlos.

PANSZA, PÉREZ Y MORÁN. Operatividad de la Didáctica. Tomo 2. México, Gemika, 1986. pp. 97-137

EVALUACIÓN SUMATIVA COMO REUNIÓN DE EVIDENCIAS DE APRENDIZAJE (FICHA 6)

La Evaluación Sumativa se ha manejado tradicionalmente en términos de promedio más que como la reunión de evidencias de aprendizaje y un "momento de recuperación de los mismos", como propone Zarzar Charur (1982).

1. Elaborar un formato para registrar y controlar la evaluación sumativa mensual y anualmente. En él se deben incluir los aspectos que con mayor frecuencia considera en la evaluación y los conocimientos y las habilidades que juzga trascendentes para la formación del niño. Por ejemplo, si usted estima que la asistencia, las tareas, la participación y la limpieza son aspectos que se deben tomar en cuenta, inclúyalos.
2. Diseñar primero un formato mensual y luego el anual. Ejemplos:

NOMBRES	ASISTENCIA	TAREAS	PARTICIPACIÓN NO / OCAS / CONT	EXAMEN INDIVIDUAL	EXAMEN EN EQUIPO BIEN / REGULAR / MAL	HABILIDADES SI / NO
---------	------------	--------	-----------------------------------	-------------------	------------------------------------------	------------------------

Ocas=ocasional, Cont=continuo, Sí y No correspondería a si han desarrollado la habilidad.

NOMBRES	ENSAYOS 123456*	ORDEN B/R/M	LIMPIEZA B/R/M	PLAN DE TRAB. DIAR D/S/NH	EX. LIBRO ABIERTO B/R/M	HABILIDADES SI/NO
---------	--------------------	----------------	-------------------	------------------------------	----------------------------	----------------------

*Los números correspondientes a cada ensayo que se elabore, se anota solamente presencia o ausencia.

Ensayos=Número entregado, B=Buena R=Regular M=Mala, D=con detalles S=someramente NH=no lo hace.

NOMBRES	TRABAJO EQUIPO B/R/M	TRABAJO INDIVID. B/R/M	TRABAJO EXTRA 12345*	HABILIDADES A DESARROLLAR AUTOESTIMA EXPRESIÓN
---------	-------------------------	---------------------------	-------------------------	---------------------------------------------------

B=Bueno R=Regular M=Mala *Número de trabajos extra que el alumno entrega

Puede observarse que cada uno de los ejemplos anteriores es distinto, habría tantos como momentos y grupos, lo importante es que usted capte la idea de que la evaluación sumativa es reunión de evidencias y no promedio. Ahora bien, ¿cómo registrar?

Se recomienda registrar por ausencia o presencia de rasgos y/o por número total de eventos, ejemplo:

NOMBRES	ASISTENCIA	TAREAS	EJERCICIOS	PLAN DE TRABAJO DIARIO	EXAMEN INDIVIDUAL	EXAMEN GRUPAL
SAÚL	23/24	5/7	6/6	23/24	11/25	10/10

En el ejemplo anterior se puede notar que el desempeño general de SAÚL fue constante y bueno, dado que tuvo 23 de 24 asistencias, hizo 5 de 7 tareas, hizo todos los ejercicios, cumplió con 23 de sus 24 planes de trabajo individuales diarios, y obtuvo calificación de 10 en el examen grupal (suponiendo que el puntaje máximo haya sido 10). Y si bien en un examen individual salió mal (obtuvo sólo 11 puntos de 25 posibles), podría deberse a múltiples circunstancias (que el día del examen estaba enfermo o no se preparó o no entendió, etc.). No obstante SU DESEMPEÑO indica que el niño se va desarrollando adecuadamente, lo cual daría elementos de juicio al profesor para determinar que SAÚL ha acreditado esta etapa de aprendizaje.

3. Elaborar el formato anual, para lo cual se procede a efectuar un análisis de los rasgos existentes en los formatos mensuales, a fin de agruparlos en el proceso de desarrollo de habilidades que el conjunto de rasgos ayudan a generar. Ejemplo:

ASISTENCIA, TAREAS Y EJERCICIOS	son evidencia de CONSTANCIA
PLANES DE TRABAJO DIARIO	son evidencia de PLANIFICACIÓN Y PREVISIÓN
LOS EXÁMENES	son evidencia de INFORMACIÓN

4. Armar el formato incluyendo los rasgos definidos y la calidad de los mismos. Ejemplo:

NOMBRES	CONSTANCIA			PLANIFICACIÓN			INFORMACIÓN		
	SI	EN DESARROLLO	NO	SI	EN DESARROLLO	NO	SI	EN DESARROLLO	NO

La estimación de los rasgos en los formatos mensuales permitirá valorar si las habilidades se han desarrollado suficientemente en cada niño.

5. La Evaluación Sumativa implica un análisis de los avances logrados, los problemas y los obstáculos. Por tanto, es importante que al efectuar la concentración de las evaluaciones sumativas mensuales redacte el análisis resultante, el cual deberá ser entregado al profesor que atienda el grupo el siguiente año.

El análisis podría resumirse atendiendo a los siguientes aspectos:

NOMBRES	HABILIDADES DESARROLLADAS	HABILIDADES EN DESARROLLO	PROBLEMAS PENDIENTES
---------	---------------------------	---------------------------	----------------------

6. Hemos destacado hasta aquí el proceso de concentración, análisis y registro de la evaluación sumativa, pero es necesario también efectuar una actividad de evaluación sumativa con los niños, para lo cual se recomienda destinar una de las últimas semanas del año para realizar una observación detallada de cómo los alumnos se desenvuelven en cada uno de los procesos que se trabajaron durante el año. Es conveniente utilizar un formato similar al propuesto para la evaluación diagnóstica (ver Ficha No.4); agregando una columna para identificar si se cumplió o no lo programado.

La evaluación sumativa "...contempla tanto los aprendizajes como el proceso seguido... no se identifica con la prueba o el examen final... Para el profesor-coordinador, es éste un momento muy importante, fuente de aprendizajes para él, ya que a través de la retroalimentación que reciba del grupo podrá evaluar el método de trabajo seguido..." (ZARZAR, 1983, p. 45) "...es una tarea que ayuda a la revisión del proceso grupal, en términos de las condiciones en que se desarrolló, los aprendizajes alcanzados, los no alcanzados, así como las causas que posibilitaron o imposibilitaron la consecución de las metas propuestas". (MORÁN, 1981, p.24).

MORÁN OVIEDO, PORFIRIO. "La Evaluación de los Aprendizajes y sus Implicaciones Educativas y Sociales" en Perfiles Educativos No. 13 México, CISE/UNAM, 1981. pp. 21-36

ZARZAR CHARUR, CARLOS. "Diseño de Estrategias de Aprendizaje Grupal. Una Experiencia de Trabajo" en Perfiles Educativos No. 1. Nueva Época. México, CISE/UNAM, 1983. pp. 34-46

EVALUACION GRUPAL (FICHA 7)

Los maestros de primaria conocen el valor del trabajo en equipo en especial con los grupos numerosos, el grupo "...es un proceso en marcha... se constituye y pasa por diversas etapas durante su existencia..." (SANTOYO, 1981, p.4) que para ser aprovechado cabalmente es organizado en pequeños equipos, interactuantes todos y generadores de la vida del grupo.

Una de las actividades que pueden realizarse por el grupo mismo es la evaluación, su evaluación.

En este caso, al igual que en la autoevaluación, su manejo requiere de un proceso previo de aprendizaje por parte de los niños que contempla tres momentos: que aprendan a trabajar y a organizarse en equipo, que aprendan a evaluar y que evalúen.

1. Trabajar y organizarse en equipo

Formar los equipos y animarlos para que se organicen (desde ponerle un nombre, color, símbolo, lema al equipo, hasta repartir funciones en él: jefe, revisor de higiene, tareas, etc.)

Vigilar que funcionen como equipo antes de enseñarlos a evaluarse.

2. Aprender a evaluar

Recapitular (en grupo) actividades realizadas: pregunte al final de cada actividad que realice: ¿Qué hicimos? Cuando hayan aprendido a recapitular por actividad, recapitule por día.

• Reflexionar lo hecho y cómo se hizo: ayúdeles formulando preguntas y escribiéndolas en el pizarrón.

¿QUÉ HICIMOS?	¿QUÉ NO ENTENDIMOS?	¿TERMINAMOS?	¿QUE FALTÓ?
---------------	---------------------	--------------	-------------

Procure que en las primeras semanas las preguntas sean pocas y siempre las mismas, hasta que los niños entiendan el propósito y el mecanismo de la evaluación.

3. Después de algunas semanas, cuando hayan aprendido el mecanismo, anote el cuadro en el pizarrón y organice a los niños en equipos para que lo llenen, agregue al cuadro la pregunta ¿POR QUÉ?

¿QUÉ HICIMOS?	¿QUÉ NO ENTENDIMOS?	¿TERMINAMOS?	¿QUÉ FALTÓ?	¿POR QUÉ?
---------------	---------------------	--------------	-------------	-----------

4: Insistir en que argumenten por qué resultó de una u otra manera una actividad.

5. Una vez que saben presentar argumentos, solicíteles que otorguen una calificación al trabajo del equipo, para lo cual presente en el pizarrón algunos criterios para que ellos los reflexionen, discutan en el equipo, asignen calificación y argumenten su decisión. Deben presentarla por escrito. Algunos ejemplos de criterios para evaluar el trabajo grupal son los siguientes:

	SI	NO	¿POR QUÉ?
¿CONCLUYERON SU TRABAJO?			
¿TODOS ENTENDIERON?			
¿AYUDARON A SUS COMPAÑEROS?			
¿CUÁNDO NO ENTENDIERON?			
¿TRABAJARON Y COLABORARON TODOS?			
¿OBTUVIERON BUENOS RESULTADOS?			
ETC.			

6. Cuando hayan entendido el procedimiento para evaluar, realice la evaluación periódicamente.

"Para que el grupo produzca intelectualmente y avance hacia el logro de sus objetivos, es conveniente que exista un clima que propicie el aprendizaje; un ambiente de libertad para pensar, expresarse, intercambiar experiencias, hacer proposiciones, señalar coincidencias, ejercer el análisis y la crítica... el conocimiento no aparece como algo terminado, sino más bien como una elaboración conjunta que parte de 'situaciones problema', en donde se elaboran hipótesis, se definen conceptos, se analizan los elementos involucrados en una situación, se proponen alternativas, se identifican los medios, se evalúan resultados..."(SANTOYO, 1981. pp. 5 y 7). "A partir de una didáctica grupal, se pueden establecer algunas condiciones para que los estudiantes puedan responsabilizarse de la asignación de notas. Esto implica generar en ellos un proceso de reaprendizaje de su papel en la escuela, propiciando que entren en un proceso de autoobservación... a reconocer lo que se va logrando... y la cantidad de problemas, contradicciones, lagunas y dudas que una nueva información plantea". (DÍAZ, 1985. p.141)

SANTOYO, RAFAEL. "Algunas reflexiones sobre la Coordinación de Grupos de Aprendizaje en Partes Educativas No.11. México, CISE/UNAM. 1981. pp. 3-19

DÍAZ BARRIGA, ÁNGEL. Didáctica y Currículum. México, Nuevo Mar, 1985. p. 150 (Col. "Problemas Educativos")

LA AUTOEVALUACIÓN EN LOS NIÑOS

(FICHAS)

Mucho se ha sugerido a los profesores que utilicen la autoevaluación con los niños. Sin embargo, pocas veces se lleva a cabo dado que se piensa que los niños no tienen capacidad y seriedad para determinar si han aprendido o no. Es muy cierto que no saben muchas cosas pero también lo es que cuando han aprendido algo YA SABEN QUE LO SABEN.

Aprender a autoevaluarse es una meta a largo plazo porque implica un proceso de aprendizaje al que ni los maestros ni los niños están acostumbrados. Se sugiere que, antes de aplicar la autoevaluación como un instrumento para la evaluación y la posterior asignación de notas o calificaciones, realice con los niños prácticas de autoevaluación con la mayor parte de las actividades que realizan.

1. Prácticas de autoevaluación.

- Al terminar cualquier actividad de las que realiza cotidianamente, acostumbre a los niños a emplear un tiempo (5 ó 10 minutos) para reflexionar sobre lo que hicieron. Hágalo primero en grupo y después en equipos.
- Recapitule las actividades que realizaron: ¿Qué hicimos? Cuando sepan recapitular, hágalo al final del día: ¿Qué hicimos hoy?

Estimule la participación de todos.

Anote en el pizarrón un cuadro sencillo para que ellos lo llenen, en el que contesten algunas preguntas que los hagan reflexionar, por ejemplo:

¿QUÉ HICIMOS?	¿QUÉ NO ENTENDIMOS?	¿TERMINAMOS?	¿QUÉ FALTÓ?
---------------	---------------------	--------------	-------------

Procure que las primeras semanas las preguntas sean pocas y siempre las mismas, hasta que los niños entiendan el propósito y el mecanismo de la evaluación.

- Después de algunas semanas, cuando hayan aprendido el mecanismo, anote el cuadro en el pizarrón y organice a los niños en equipos para que llenen uno por equipo, después de discutirlo, agregue al cuadro la pregunta ¿POR QUÉ? (ver Ficha No.7).

¿QUÉ ENTENDIMOS?	¿TERMINAMOS?	¿QUÉ FALTÓ?	¿POR QUÉ?
------------------	--------------	-------------	-----------

- Cuando hayan aprendido a evaluar en grupo y en equipo es el momento para que asignen notas. Empiece por solicitarles asignación de notas individualmente y luego por equipo. Para ello enséñeles a utilizar algunos criterios, anótelos siempre en el pizarrón y pídale que entreguen su evaluación por escrito. Ejemplo:

En la actividad que realizamos:

	SI	NO	¿POR QUÉ?
¿APRENDISTE?			
¿TRABAJASTE CON EMPENO?			
¿HICISTE CON CUIDADO LAS COSAS?			
¿REDACTASTE TU APUNTE?			
¿TUVISTE RESULTADOS CORRECTOS?			
etc.			

¿QUÉ CALIFICACIÓN TE PONDRÍAS?

Es indudable que las preguntas y su cantidad pueden variar en función del contenido y/o proceso que se evalúa. Usted mejor que nadie puede determinar dichas preguntas.

- Finalmente, enséñeles a los niños que la calificación que se asigna y/o asignan a otros compañeros.

será tomada en cuenta para su calificación final.

NOTA: Con los niños de primero y segundo grados esta estrategia sólo se puede llevar a cabo oralmente y sólo con respecto a la evaluación, la acreditación no es asequible para ellos, por tanto es el profesor quien puede registrar los resultados.

"No podemos dejar de mencionar las múltiples dificultades que entraña la experiencia en sí misma, dado que el alumno está generalmente viciado respecto a la capacidad de auto-observación y de actitud crítica consigo mismo; sin embargo, podemos afirmar que en algunas de estas experiencias los alumnos han podido manifestarse críticamente frente al conocimiento, frente a su aprendizaje y frente a su calificación, manifestación que es elaborada en presencia del grupo, que en este momento permite al sujeto reflexionar en voz alta sobre sí mismo. Lejos de lo que se puede pensar, en estas experiencias la generalidad de los alumnos no obtiene notas altas, y nuestro problema, como coordinadores, ha consistido en ayudar al grupo a pensar sobre su aprendizaje y evitar la complicidad que se puede generar por no reflexionar sobre los elementos de aprendizaje que se ha vivido". (DÍAZ, 1985. p.142).

Es importante señalar que Díaz Barriga refiere esta afirmación a la evaluación grupal y además señala que es una experiencia que está a prueba. En este documento se ha tomado como referente para la auto-evaluación dado que la estrategia se presenta combinada.

DÍAZ BARRIGA, ÁNGEL. Didáctica y Currículum. México, Nuevo Mar, 1985. p.150 (Colección "Problemas Educativos").

EVALUACIÓN DE LOS CONTENIDOS

(FICHA 9)

La evaluación de la información (o contenidos) que los alumnos han logrado aprehender o hacer suya es una de las actividades que usualmente se realizan en la primaria. Es, de hecho, casi la única actividad que se evalúa o, más bien, se mide. Sin embargo, como ya se ha señalado en otras fichas es sólo una de las actividades que se debe evaluar pues existen otras cuya evaluación es insoslayable, como lo son los procesos de desarrollo de los alumnos, el desempeño del grupo y el desempeño del profesor.

Existe en los profesores de primaria una amplia experiencia en la medición de los contenidos. Se ha privilegiado tradicionalmente la utilización del examen escrito en sus distintas modalidades: opción múltiple, canevá, falso-verdadero, correlación, ordenación y/o secuencia de eventos, complementación, etc. En esta ficha se proponen cinco modalidades que al ser combinadas permiten evaluar los contenidos. Dichas modalidades han sido hasta ahora poco utilizadas en la Educación Primaria: examen a libro abierto, examen en equipo, examen global mensual para resolver en casa, trabajo de composición e investigación.

Examen a libro abierto

1. Diseñe las preguntas del examen, se sugiere que sean preguntas de respuesta cerrada, corta y única para los niños de primero y segundo grados; y preguntas de respuesta abierta, problematizadora y amplia para los niños de tercero a sexto grados. Vigile que la cantidad de preguntas y su nivel de dificultad sea acorde tanto con lo que han estudiado como con sus procesos de desarrollo.

2. Haga una lista de preguntas amplia, de tal forma que se pueda fraccionar y organizar en varios cuestionarios, de acuerdo con las características y adelanto de los niños. Por ejemplo: Se puede hacer la misma pregunta de manera problematizadora y abstracta para alumnos con gran capacidad de análisis; una pregunta cerrada y concreta para alumnos que aprenden más lentamente, etc.

Recuerde que la evaluación debe ser un momento más de aprendizaje.

Una forma práctica de elaborar preguntas y organizarlas, sin que ello ocupe mucho tiempo, es redactarlas en el momento que se está impartiendo el contenido. Tenga a mano siempre una provisión de tarjetas y anote la pregunta, por lo menos en idea general.

Después revísela y corrija la redacción si es necesario; en la misma tarjeta anote el título del contenido, la cita bibliográfica de los libros de texto del niño donde se puede ubicar la respuesta y la fecha de elaboración. Las mismas tarjetas pueden servirle para repartirlas a los niños el día del examen (sin bibliografía).

3. Avise a los niños con anticipación del día que habrá examen y enfatice la necesidad de llevar los libros correspondientes.

4. Enséñeles y acostúmbrelos a incluir en su respuesta la referencia bibliográfica (incluidas las páginas) donde la encontraron; es de particular importancia que aprendan a buscar, seleccionar y citar información.

5. Pídale (y hágalo hasta que se acostumbren) escribir en su respuesta, además de lo expresado en el libro o libros consultados, su opinión personal.

6. Si decide asignar calificación al examen, es conveniente que fije puntuación por cada uno de los siguientes elementos: opinión del autor, referencia bibliográfica y opinión personal.

7. El día del examen presente el cuestionario, haga que sea leído y aclare las dudas que se susciten; explique claramente cuál es la forma de resolver el examen (pueden utilizar los libros y apuntes que crean necesarios).

Examen en equipo

1. Elabore el cuestionario de examen de la misma manera que en el caso anterior.

2. Organice equipos de trabajo y explique las normas del examen.

a. No podrá haber error alguno puesto que son varias personas pensando.

b. Deberán entregar un solo examen escrito, pero en el que todos deben intervenir y anotar al final.

del párrafo que les haya correspondido, su nombre.

c. El texto deberá entregarse en limpio y sin errores ortográficos.

d. Todas las respuestas deberán incluir la referencia bibliográfica del libro o los libros que utilizaron.

e. Todas las respuestas deberán incluir, además de la opinión del autor, un ejemplo y una conclusión de equipo.

3. Asigne tiempo suficiente para que los niños puedan utilizar sus libros y apuntes, discutir entre ellos, llegar a acuerdos y redactar.

Se asombrará especialmente observar cómo los niveles de exigencia entre los niños para su trabajo son mucho más altos y productivos que en otros tipos de trabajo:

Examen global mensual para resolver en casa

1. Elabore el cuestionario de examen de la misma manera que en los casos anteriores.

2. Este examen tiene la particularidad de que deberá ser entregado a los niños antes de empezar la enseñanza del contenido.

Asegúrese de que lo lean y planteen todas las dudas que tengan sobre el cuestionario.

3. Explíqueles que cada respuesta debe incluir: opinión del autor, referencia bibliográfica, ejemplo y conclusión.

4. Explíqueles que deberán resolver el cuestionario diariamente cuando se vayan tratando los diversos temas. Cada vez que termine una clase, recuérdelos que en el transcurso de la clase se dio respuesta a "tales preguntas" del cuestionario; solicite su participación para que digan cuáles fueron las respuestas. Indíqueles que por la tarde deberán redactar sus respuestas. Acostúmbrelos a anotar algún símbolo convencional que les indique que ahí está una respuesta de examen en sus apuntes y en los libros.

5. Empezar por aplicar exámenes de un día, luego semanales y luego mensuales. Verá cómo los niños se acostumbran y poco a poco van aprendiendo a registrar y organizar su información.

Trabajos de composición e investigación

Este instrumento se recomienda para los alumnos de cuarto a sexto grados.

Para desarrollarlo es indispensable que los niños aprendan primero cómo se hace una composición y una investigación.

Una vez que han madurado y entendido los procedimientos podrán aplicarse para la evaluación.

1. Asigne el tema de composición por medio de un problema a resolver.
2. Señale los requisitos del trabajo (estructura, presentación, contenido, etc.).
3. Indique bibliografía y asigne un tiempo prudente para su elaboración.
4. Indique criterios de evaluación.
5. Revise periódicamente el avance de los niños, recuerde que ellos aún están conformando el hábito de la disciplina para el trabajo. Estimúelos y oriéntelos.

"Somos conscientes que este tipo de examen exige una concepción distinta del proceso de enseñanza-aprendizaje, pero tendría a su favor practicar una evaluación cualitativa capaz de canalizar creatividad, interpretación personal y/o grupal, juicio crítico y material bibliográfico, etc... La calidad de estas pruebas depende de su elaboración: no pueden ser improvisadas; su construcción exige pensar detenidamente en las cuestiones que estimulen conductas que impliquen la capacidad de manejo de la información, cómo sugerir nuevas salidas a un planteamiento, aplicar principios, leyes, tendencias, normas o nuevas situaciones; fundamentar aprendizajes que impliquen procesos mentales superiores, tales como la capacidad para pensar, organizar y aplicar la información recibida; integrar aprendizajes; la capacidad para producir, organizar y expresar ideas o crear formas originales, etc.". (PANSZA, PÉREZ Y MORÁN, 1986. p.137)

PANSZA, PÉREZ Y MORÁN. Operatividad de la Didáctica. Tomo 2. México, Gemika, 1986. p. 137

LA ACREDITACIÓN DE LOS APRENDIZAJES

La acreditación de los aprendizajes se ha asumido como asignación de notas o calificaciones, exclusivamente, lo cual ha generado una visión de la evaluación que apoya la noción de aprendizaje como producto y no como proceso. Así, el profesor se ha acostumbrado a calificar en términos numéricos ciertos resultados que presentan los niños (exámenes, resultados de ejercicios, trabajos, etc.), pero ha olvidado sistematizar la actividad que cotidianamente realiza: estimar "cómo van los niños", si avanzan o no en su formación más que en los conocimientos que poseen.

Para registrar la acreditación de los aprendizajes se sugiere un procedimiento que combina algunas de las estrategias sugeridas a lo largo de este documento; por tanto debe revisar las fichas 6, 7 y 8.

1. La acreditación del aprendizaje podría realizarse combinando tres instancias de juicio:

EL PROFESOR

EL EQUIPO DE TRABAJO

EL ALUMNO

2. Para el profesor, la manera de asignar la calificación deriva de la reunión de evidencias de aprendizaje que realiza a lo largo del proceso, es decir, si recuerda lo planteado en la Ficha 6 sobre evaluación sumativa.

NOMBRES	ASISTENCIA	TAREAS	EJERCICIOS	PLAN DE TRABAJO DIARIO	EXAMEN INDIVIDUAL	EXAMEN GRUPAL
SAÚL	23/24	5/7	6/6	23/24	11/25	10/10

Los elementos ahí verificados le permitirán establecer una calificación justa para el alumno, que no derive de un solo elemento, como el examen por ejemplo, sino que tome en cuenta el desempeño general del alumno. En el caso que se ejemplifica, si la acreditación se tomara a partir del examen individual, la calificación de Saúl sería de 4 (11 aciertos entre 25 posibles), por lo tanto reprobatoria.

No obstante, su desempeño en el resto de los trabajos ha sido excelente, de tal forma que, considerándolo, su calificación debería ser mayor, quizá entre 8 y 9. Es el profesor quien debe tomar esta decisión con base en su conocimiento del niño.

3. El equipo de trabajo podría aprender a evaluar a sus compañeros si el maestro les enseña. Para ello se sugiere:

- Realizar un ejercicio diario de evaluación (ver Fichas 7 y 8).
- Utilizar la misma rutina hasta que los niños entiendan los aspectos a evaluar y el procedimiento.
- En principio, solamente, se les pediría que evaluaran en términos cualitativos diciendo si lo hicieron bien o mal, rápido o despacio, si colaboraron o no.
- Cuando hayan aprendido a expresar sus juicios se les pediría que asignaran calificación.

4. Después de una temporada en que los niños utilicen la evaluación en equipo, podrá pedirles que inicien prácticas de autoevaluación (ver Ficha 8), hasta que sean capaces de emitir una nota para sí mismos.

5. Finalmente, el promedio entre las tres calificaciones emitidas, podría ser la calificación definitiva.

Si bien la acreditación en "estricto" se refiere al hecho de avalar o no si un alumno ha aprobado un curso

y la calificación, aluden a la síntesis y expresión de la acreditación en una calificación o nota, en el documento se asumen como un procedimiento único. Es importante recordar que "Se puede calificar a los alumnos sin evaluar el curso ni los aprendizajes [como generalmente sucede]. Así como también se pueden evaluar éstos sin necesidad de adjudicar calificaciones" (ZARZAR, 1983, p. 45)

"La distinción entre evaluación y acreditación da pauta a la existencia de un momento de evaluación en el que el sujeto y el grupo reflexionan sobre su proceso de aprendizaje. Indudablemente, efectuar la evaluación es un paso necesario para poder dejar en manos del grupo la acreditación... Si en la evaluación grupal los coordinadores perciben que el nivel alcanzado por proceso del grupo les permite responsabilizarse de su acreditación, se requiere llevarla adelante. Es aquí y no al principio del curso cuando se puede decidir si el grupo está en posibilidades de autoasignarse una nota". (DÍAZ, 1985, p.143)

ZARZAR CHARUR, CARLOS. "Diseño de Estrategias para el Aprendizaje Grupal", en Perfiles Educativos No.1. Nueva Época, México, CISE/UNAM, 1983, pp. 34-46.
DÍAZ BARRIGA, ÁNGEL, Didáctica y Currículum. México, Noved Mar, 1985 p.150 (Col. "Problemas Educativos").

LA LISTA DE COTEJO COMO UNA FORMA DE REGISTRAR OBSERVACIONES PARA EVALUAR (FICHEA III)

El maestro, por su experiencia en el trato con los niños, conoce fases o etapas por las que pasan los niños. Sabe bien, por ejemplo, que cuando un niño va a iniciar su aprendizaje de la lecto-escritura, "hace que lee", escribe palabras incompletas ("taja" por tinaja). Ambos son pasos de maduración para llegar a leer-escribir, por tanto están correctos, de acuerdo con la madurez del niño. Es precisamente el maestro quien mejor conoce esos procesos de desarrollo. Se considera que es él quien mejor puede elaborar registros de observación para la evaluación de sus alumnos: Hasta ahora, dichos registros no se acostumbra en la escuela primaria; no obstante, se sugiere su utilización para la evaluación en el aula.

1. Elija la o las habilidades que pretende evaluar.
2. Liste los rasgos que va a observar, son las etapas, fases o momentos por los que usted considera pasan los niños para avanzar en su proceso de desarrollo de habilidades.
3. Elabore la lista de cotejo. Por lo general estas listas presentan tres-elementos estructurales:

- LOS DATOS DE IDENTIFICACIÓN: Título, Fecha, Alumnos observados y Observador
- LOS RASGOS A OBSERVAR: Por lo general en columna ubicada a la izquierda
- LOS ESPACIOS PARA EL REGISTRO: Se ubican en columna a la derecha

Se le sugiere que utilice los espacios para el registro no sólo para ausencia o presencia de rasgos, sino para alguna clasificación del grado o nivel que usted juzgue conveniente. Aunque es importante tener presente que ésto implicaría una combinación del instrumento con Escala de Actitudes (ver Anexo 2, Ejemplo de Lista de Cotejo).

4. No pretenda diseñar todas las listas de cotejo que necesita: elabore, use, pruebe y corrija una por vez y vaya integrando poco a poco su Batería de listas que le serán útiles en el futuro, en especial si trabaja con el mismo grado. Evaluar por desarrollo de habilidades implica una forma distinta de trabajo; por tanto, la utilización de listas de cotejo para tal efecto debe ser paulatina. Puede decirse con seguridad que significará un aprendizaje para el maestro, y en consecuencia, llevará un tiempo largo: dos o tres años.

5. Revise los rasgos que listó y asegúrese de que describen adecuadamente el proceso a observar.

La lista de cotejo es un instrumento apropiado para aplicar la técnica de observación "...se usan para evaluar procesos, productos y aspectos de desarrollo social, donde la evaluación de las características puede limitarse a juicios de 'presente-pasado', sí-no, etc.". (GRONLUND, s/f).

"Quizá el mayor valor de las técnicas de observación sea el que hacen posible obtener la información del comportamiento tal como ocurre". (SELLTIZ, 1974. p. 229).

Para el maestro observar el comportamiento de sus alumnos es la clave en la determinación de la madurez de los procesos de desarrollo.

GRONLUND, N.E. "Técnicas de Observación", en Cuadernos de Tecnología Educativa, Taller de Evaluación del Aprendizaje No. 2. México, CEUTES/UNAM, s/f. pp. 57-64

SELLTIZ, et al. "Métodos de Observación", en Métodos de Investigación en las Relaciones Sociales. Madrid, Rialp, 1974. pp. 228-266

LA OBSERVACIÓN COMO EL MEDIO FUNDAMENTAL PARA LA EVALUACIÓN

(FICHA 12)

La observación es una de las actividades cotidianas que los profesores de primaria realizan con mayor asiduidad, sin embargo las conclusiones que extraen, por lo general, no son sistematizadas. Es necesario que los profesores asuman la importancia de sistematizar científicamente la observación en particular, si se considera que es un elemento clave para la evaluación.

1. Determine qué instrumento de registro va a utilizar para la observación (Diario de Campo, Escala de Actitudes o Lista de Cotejo).
2. Defina y delimite el objeto de observación (actividades, niños, documentos, etc.).

3. Establezca el (los) objetivo(s) de la observación.
4. Seleccione y defina las categorías de análisis y los rasgos de cada categoría (Por ejemplo: Categoría=Participación; Rasgos= participación en: juegos, trabajo en equipo, tareas, iniciativas, etc.).
5. Diseñe el formato de registro que debe incluir: Datos de identificación, Objetivo, Rasgos a Observar y Espacios para el Registro. Incluya instrucciones para el llenado, si el observador no va a ser usted.
6. Realice uno o dos ejercicios de observación para corregir el instrumento.
7. Establezca por escrito la forma de Interpretar (vaciado, concentración y análisis) los datos recabados por el Instrumento antes de utilizarlo definitivamente.
8. Realice la observación.

"La observación no es solamente una de las más útiles actividades de la vida diaria; es un instrumento primordial de la investigación científica. La observación se convierte en técnica científica en la medida que: 1) sirve a un objetivo ya formulado...; 2) es planificada sistemáticamente; 3) es controlada... y relacionada con proposiciones más generales [teoría]... La observación puede tener lugar en situaciones auténticas de la vida ordinaria o bien en el laboratorio. Los procedimientos de observación pueden ir, desde la flexibilidad total, guiada solamente por la formulación del problema a ser estudiado y algunas ideas generales acerca de los aspectos de importancia probable, hasta el uso de minuciosos instrumentos... El observador puede participar activamente en el grupo que está observando; puede ser definido como un observador que no forma parte del grupo o bien su presencia puede ser desconocida para la gente que está observando" (SELLTIZ, 1974. pp. 228-230).

Se considera que, para los profesores en general, los instrumentos idóneos para efectuar la observación son el diario de campo, la escala de actitudes y la lista de cotejo.

El Diario de Campo: "... es un instrumento de recopilación de datos, con cierto sentido íntimo recuperado por la misma palabra diario, que implica la descripción detallada de acontecimientos, y se basa en la observación directa de la realidad, por eso se denomina "de campo". En el caso del trabajo docente, el diario de campo es instrumento de reflexión y análisis del trabajo en el aula, y por esto mismo un trabajo de descripción, valoración y explicación de los niveles de significación de la práctica educativa... requiere del estudio teórico y práctico de los comportamientos... difiere de la observación ingenua [porque] cuenta con un apoyo teórico en la problematización de las dimensiones del trabajo docente... Problematizar, pues, supone plantearse interrogantes sobre el espacio y el tiempo en que se desarrolla el proceso enseñanza-aprendizaje..." (GERSON, 1979. pp. 5 y 6).

Escala de Actitudes: "Las escalas de evaluación simple... se usan a menudo para registrar observa-

ciones cuantificadas de una situación social. Puedo usárselas para describir la conducta de individuos, los cambios en la situación que los rodea y muchos otros tipos de datos... La expresión 'escala simple' denota una escala con un conjunto de puntos que describen diversos grados de la dimensión observada... Las escalas de evaluación se usan principalmente en dos formas: 1) para registrar la conducta a intervalos frecuentes a lo largo de un proceso de interacción social o 2) para evaluar la naturaleza de un hecho social completo después de terminado... Si no se está seguro acerca de cuáles son los indicios de un tipo dado de conducta, las escalas pueden ser útiles para definirlos". (HEYNS Y ZANDER en FESTINGER Y KATZ, 1987. pp. 363-367)

HEYNS y ZANDER. "Observación de la Conducta de Grupo", en FESTINGER y KATZ. *Los Métodos de la Investigación en las Ciencias Sociales* (tr. Eduardo Masulio), 1a. reimp. México, Paidós. 1987. pp. 353-385 (Col. "Studio/Básica")

GERSON, BORIS. "Observación Participante y Diario de Campo en el Trabajo Docente" en *Perfiles Educativos* No. 5. México, CISE/UNAM, 1979. pp. 3-22

SELLTIZ, y otros. "Métodos de Observación" en *Métodos de Investigación en las Relaciones Sociales*, 7a. ed. Madrid, Rialp, 1974. pp. 228-266. Citado en SEP. *Metodología para la Evaluación Educativa*. México, SEP/DGEE/DEPE, 1987. pp. 1-20 (Curso Sobre Evaluación del Proceso Educativo, Tomo 2, Vol. II)

ANEXOS

ANEXO I. ANÁLISIS DE PROGRAMAS

CONTENIDOS Y HABILIDADES

3er. grado
Español

Lectura/exploración de diversos materiales escritos/contestación de materiales escritos/lectura comprensiva/identificación de elementos de la lectura/reformulación de textos/creación de textos/formación de acervos (antologías, colecciones, bibliotecas, etc.).

Investigación/exploración del diccionario, el periódico y otras fuentes de consulta/identificación de las convenciones para el manejo de cada fuente de consulta/localización de información.

Escritura/elaboración de textos con tema prescrito/elaboración de textos con tema libre/elaboración de textos de uso cotidiano (recados, avisos, cartas, etc)/elaboración de textos escolares e Informativos/juegos sintácticos con enunciados.

Expresión oral/diálogo/exposición de anécdotas/exposición de temas/debates.

CONTENIDOS Y HABILIDADES

3er. grado
Matemáticas.

Sistema decimal/agrupamientos/representación posicional/conteo.

Números naturales/antecesor y sucesor/operaciones.

Problemas/planteamiento/resolución.

Medición/unidades de longitud/unidades de área/estimación/unidades de capacidad/unidades de peso/unidades de tiempo/uso de balanza/lectura de reloj.

Espacialidad/ubicación/representación de ubicaciones/objetos tridimensionales.

Figuras/análisis de figuras/descomposición de figuras/interpretación de croquis/construcción de sólidos geométricos.

Tratamiento de la información/recolección de datos/representación.

ANEXO 3 PLAN INDIVIDUAL DE RECUPERACION MENSUAL

NOMBRE: _____

FECHA: _____

1. ANOTA EN LISTA CUÁLES DE LAS COSAS QUE HICIMOS EN LA SEMANA DE EVALUACION NO SUPISTE O TE SALIERON MAL.

- a. NO SUPE DIVIDIR FRACCIONES.
- b. SE ME OLVIDA EMPEZAR CON MAYÚSCULAS CUANDO HAY PUNTO Y APARTE.
- c. SE ME OLVIDA LAVARME LAS MANOS DESPUÉS DE JUGAR

2. ESCRIBE QUE COSAS VAS A HACER PARA MEJORAR LO QUE ANOTASTE EN LA LISTA.

- a. PEDIRLE A LA MAESTRA EJERCICIOS EXTRA PARA HACER EN MI CASA CON AYUDA DE MAMÁ.
- b. VOY A ESCRIBIR CON ROJO LAS MAYÚSCULAS AL EMPEZAR UN PÁRRAFO.
- c. LE VOY A ENSEÑAR A MI HERMANO PEQUEÑO A LAVARSE LAS MANOS.

A LOS PROFESORES DE GRUPO

NIVEL EDUCATIVO _____

ESTRATO _____

El objetivo principal de este documento es proporcionar al docente de educación básica algunas sugerencias útiles para apoyar la práctica de la evaluación en su salón de clase, por esta razón la opinión de los destinatarios de este material tiene singular importancia para continuar desarrollando trabajos como éste. Para conocer estas opiniones, la DGE ha formulado las preguntas que aparecen a continuación; y le agradecerá el envío de este cuestionario resuelto a las oficinas del CETE-Veracruz, a través de la Jefatura de Sector y/o de la Supervisión Escolar correspondiente. Si el espacio no es suficiente, por favor use hojas adicionales.

1. ¿Las estrategias sugeridas le parecen claras?

Sí () No ()

¿Por qué?

2. Las situaciones o problemas de evaluación que se establecen al principio de las fichas, ¿plantean problemas que usted ha enfrentado al realizar la evaluación en el aula?

Sí () No ()

¿Por qué?

3. ¿La bibliografía mínima le es útil como material de consulta para desarrollar la evaluación en el aula?

Sí () No ()

¿Por qué?

4. ¿Cuál(es) de las doce fichas cree usted que le resultará(n) de mayor utilidad?

Especifique (X) cuál(es): (1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12)

¿Por qué?

5. ¿El apoyo teórico a cada una de las fichas le parece claro?

Sí () No ()

¿Por qué?

6. ¿Qué le parecen, en general, los siguientes aspectos del presente documento?

	Adocundo	Regulamento Adecuado	Inadecuado
Presentación	()	()	()
Adecuación a las condiciones de su trabajo en el aula	()	()	()
Oportunidad en la distribución	()	()	()

7. ¿Cómo considera usted que podría mejorarse este documento?

BIBLIOGRAFÍA

- DÍAZ BARRIGA, ÁNGEL. "Problemas y Retos del Campo de la Evaluación Educativa", en *Perfiles Educativos* No. 37. México, CISE/UNAM, 1987. pp. 3-15
Didáctica y Currículum. México, Nuevo Mar, 1985. p. 150 (Colección Problemas Educativos)
- GERSON, BORIS. "Observación Participante y Diario de Campo en el Trabajo Docente", en *Perfiles Educativos* No. 5. México, CISE/UNAM, 1979. pp. 3-22
- GRONLUND, N.E. "Técnicas de Observación" en *Cuadernos de Tecnología Educativa. Taller de Evaluación del Aprendizaje*. No. 2. México, CETES/UNAM, s/f. pp. 57-64
- HEYNS y ZANDER. "Observación de la Conducta de Grupo" en FESTINGER y KATZ. *Los Métodos de la Investigación en las Ciencias Sociales*; (tr. Eduardo Masulio). 1a. reimp. México Paidós, 1987. pp. 353-385. (Col. "Studia Básica")
- MORÁN OVIEDO, PONTIFICIO. "La Evaluación de los Aprendizajes y sus Implicaciones Educativas y Sociales", en *Perfiles Educativos* No. 13. México, CISE/UNAM, 1981. pp. 21-36
"Propuesta de Evaluación y Acreditación en el Proceso de Enseñanza Aprendizaje, desde una Perspectiva Grupal", en *Perfiles Educativos* Nos. 27 - 28. México, CISE/UNAM, 1985. pp. 9-25
- PANSZA, PÉREZ Y MORÁN. *Operatividad de la Didáctica*. Tomo 2. México, Gomika, 1986. p. 137
- SANTOYO, RAFAEL. "Algunas Reflexiones sobre la Coordinación de Grupos de Aprendizaje", en *Perfiles Educativos* No. 11. México, CISE/UNAM, 1981. pp. 3-19
"En Tomo al Concepto de Interacción", en *Perfiles Educativos* Nos. 27-28. México, CISE/UNAM, 1985. pp. 66-71
- SELLTIZ, y otros. "Métodos de Observación", en *Métodos de Investigación en las Relaciones Sociales*, 7a. ed. Madrid, Rialp, pp. 228-266. Metodología para la Evaluación Educativa. México, SEP/DGEE/DEPE, 1987. (Curso Sobre Evaluación del Proceso Educativo. Tomo 2. Vol. II)
- SECRETARÍA DE EDUCACIÓN PÚBLICA. *Educación Básica. Primaria. Plan y Programas de Estudio 1993*. México, SEP, 1993. p. 140
- SECRETARÍA DE EDUCACIÓN PÚBLICA. *Metodología para la Evaluación Educativa. Curso Sobre Evaluación del Proceso Educativo*. Tomo 2. México, SEP/DGEE/DEPE, 1987. p. 267
- ZARZAR CHARUR, CARLOS. "Diseño de Estrategias para el Aprendizaje Grupal", en *Perfiles Educativos* No. 1. Nueva Época. México, CISE/UNAM, 1983. pp. 34-46

Tests hechos por el profesor

Ya sea que nos agrade o no, el hecho es que tenemos que medir los resultados de la enseñanza. Es verdad que los expertos en elaboración de tests no siempre realizan bien su trabajo y que cuentan con la ayuda de especialistas en las asignaturas. Hemos señalado los problemas de validez y confiabilidad en la construcción de tests estandarizados. ¿Cómo podemos esperar entonces que un maestro que sólo ha recibido un curso sobre ellas elabore buenos tests?

Es cierto que usted nunca producirá uno que realmente satisfaga las normas ideales de la construcción de tests, no importe lo mucho que leen sobre el tema en éste u otros libros. Sin embargo, probablemente sus tests serán lo bastante buenos para sus propósitos, si aplica los métodos correctos de construcción de tests. La lectura de este capítulo no resolverá todos sus problemas en este respecto, pero sí le será útil. Recuerde que los tests hechos por el profesor, del mismo modo que los estandarizados, son un reflejo de las metas y procedimientos educativos; pero un reflejo más preciso porque el profesor sabe cuáles son los objetivos específicos y qué métodos se han empleado para lograrlos.

LOS TESTS ESTANDARIZADOS EN CONTRASTE CON LOS HECHOS POR EL PROFESOR

No debe haber antagonismo entre los tests estandarizados y los hechos por el profesor. La mayoría de las escuelas usan ambos tipos de instrumentos; y cada uno tiene sus propias ventajas y desventajas. Evidentemente, aquí no se considerarán los tests de inteligencia, de personalidad, aptitud especial e intereses elaborados por el profesor. El problema es cuándo usar el test de aprovechamiento estandarizado y el de aprovechamiento hecho por el profesor. Una regla general es que el test hecho por el profesor siempre es el mejor instrumento para apreciar los resul-

KARÆL, Loid J. "Tests hechos por el profesor" en:
Medición y evaluación escolar. México, Trillas, 1986.

TABLA 25. Ventajas y limitaciones de los tests de aprovechamiento estandarizados y no estandarizados.

ESTANDARIZADOS

Criterio	Ventajas	Limitaciones
1. Validez		
a) Curricular	Selección cuidadosa hecha por personas competentes. Se ajusta a las situaciones típicas.	Inflexibles. Demasiado generales para llenar completamente requisitos locales, especialmente en situaciones desacomunadas.
b) Estadística	Con los mejores tests, alta.	Los criterios a menudo son impropios o inconfiables. La magnitud de los coeficientes dependen de la amplitud de la capacidad de los grupos examinados.
2. Confiabilidad	En los mejores tests, bastante alta, a menudo de 0.85 o más en formas similares.	La alta confiabilidad no es garantía de validez. La confiabilidad depende también de la amplitud de la capacidad del grupo examinado.
3. Uso		
a) Facilidad de administración	Procedimiento definido, tiempo fijo, claves, etcétera. Economía de tiempo.	Los manuales, a veces inadecuados, deben estudiarse detenidamente.
b) Facilidad de calificación	Reglas definidas, claves, etcétera. Muy rutinarias.	La calificación a mano puede requerir de mucho tiempo y ser monótona. Es preferible la calificación mecánica.
c) Facilidad de interpretación	Los mejores tests tienen normas adecuadas. Bases útiles de comparación. Formas equivalentes.	Las normas a menudo se confunden con los estándares. Algunas normas son defectuosas. A menudo carecen de normas para tipos de escuelas y niveles de capacidad diferentes.
Resumen	Conveniencia, comparabilidad, objetividad. Pueden obtenerse formas equivalentes.	Inflexibilidad. Quizás sean solamente aplicables a una situación en particular.

NO ESTANDARIZADOS

Ensayo	Limitaciones	Ventajas	Objetivos
Útil para clases avanzadas de inglés; facilita el aprendizaje de un idioma. Estimula los buenos hábitos de estudio.	Muestreo limitado. Es posible abarcar. Mezcla factores del lenguaje en todas las calificaciones. Generalmente es desconocidos.	Muestreo extenso de la materia. Flexible en su uso. Demuestra el alarde. Se compara favorablemente con los tests estandarizados.	Medidas de aprendizaje en el aula. Preparación para el examen.
Fácil de preparar. Fácil de aplicar.	Falta de uniformidad.	Generalmente su confiabilidad es baja.	Algunas veces se aproximan a los tests estandarizados.
			Las instrucciones son uniformes. Se economiza tiempo para el alumno.
	Lenivas, inciertas y subjetivas.	Reglas definidas, claras, etcétera. Pueden hacerse a mano o mecánicamente.	Son manuales.
No tienen normas. Significado dudoso.			No se dispone de normas al principio.
Útil cuando se aplica con muchos tests en unos cuantos campos especiales.	Muestreo limitado. Calificación subjetiva. Requiere de mucho tiempo.	Muestra amplia. Calificación objetiva. Flexibilidad.	Preparadas exigencia de destreza y tiempo.

tados de la enseñanza local. Por otra parte, el test estandarizado se usa para proporcionarle información al profesor, a la escuela y al estudiante con respecto al nivel del aprovechamiento local en comparación con las normas nacionales. En la tabla 25 se hace una comparación de las fortalezas y debilidades relativas de los tests estandarizados y los elaborados por el profesor o no estandarizados.

PROPOSITOS DE LOS TESTS HECHOS POR EL PROFESOR

"El test es el mensaje." El profesor, a través de sus tests, le dice al estudiante, en la forma más directa y significativa, lo que piensa que es realmente importante. Desafortunadamente, el resultado de esto es que a menudo el alumno estudia para un test y no para aprender. Aunque podemos desalentar el estudio encaminado exclusivamente a pasar tests, no cabe duda que será imposible eliminarlo. Por tanto, usted debe comunicar lo que piensa que es importante de la materia a través de los exámenes.

Los tests le proporcionan al profesor indicios valiosos de la efectividad de su enseñanza. El profesor puede preguntarse: "¿Están aprendiendo los estudiantes lo que quiero enseñarles?" Y después hallar la respuesta en los resultados del test.

Además, los tests le proporcionan al profesor información sobre el progreso individual. Sin estos datos, es imposible tomar decisiones relativas a reprobaciones y aprobaciones, formación de grupos, cursos de rectificación, y tampoco podría brindarse al alumno avanzado experiencias fecundas. Los tests también ayudan a evaluar al grupo en su conjunto.

Los tests, por supuesto, proporcionan solamente una base parcial para asignar calificaciones. Cuando se usan adecuadamente pueden prevenir la desilusión de los padres al recibir la boleta de calificaciones de sus hijos. El padre que ha visto los tests de su hijo sabe muy bien qué calificaciones recibirá.

Revisando los tests con los padres y los profesores se contribuye a que el niño aprenda los conocimientos que todavía no haya asimilado. El profesor podrá detectar los problemas que perturban al estudiante y proporcionarle ayuda.

PLANEACION

Ya lo ve, maestra Enriqueta, es su primer mes de trabajo y lo ha hecho bastante bien. Ahora tiene usted que decidir cómo va a evaluar

lo aprendido por los estudiantes ya que tiene que entregar la boleta de calificaciones la semana próxima. "¿Qué clase de test va a usar?"

Los tests apropiados no saltan a la vista. Se planean cuidadosamente cuando se establecen las metas globales del curso. Antes de que comience el curso, tiene que decidir el tipo de test que va a usar. La extensión del contenido que se va a examinar tanto del libro de texto como de las discusiones de clase y así por el estilo.

DEFINICION DE OBJETIVOS

La definición de objetivos no será cosa nueva para quienes hayan recibido un curso de métodos educativos. Usted sabe que la definición de metas u objetivos es componente fundamental del proceso de enseñanza. Noll (1965) lo expresa muy convincentemente: "tratar de enseñar y evaluar sin definir objetivos es como iniciar un viaje sin saber a dónde se va. Acaso sea agradable andar sin rumbo durante un tiempo, pero es muy difícil que pueda progresarse si no se tiene una meta fija", (página 104).

Objetivos generales

Comúnmente hay dos clases de objetivos generales y específicos. Por ejemplo, hace muchos años un grupo de educadores estableció los objetivos que se conocieron después como los *siete principios cardinales* de la educación (Departamento del Interior de los Estados Unidos, 1918). Estos objetivos generales son los siguientes:

1. Promover la salud
2. Enseñar a conocer los procesos fundamentales
3. Fortalecer los lazos familiares
4. Ayudar a seleccionar una ocupación
5. Ofrecer educación cívica
6. Aprovechar bien el tiempo libre
7. Estimular los valores éticos

Otro ejemplo de objetivos generales es el estudio clásico de ocho años (Aikin, 1942).

1. La adquisición de métodos eficaces de raciocinio
2. El cultivo de hábitos útiles de trabajo y estudio
3. Inculcar actitudes sociales
4. La adquisición de una gama amplia de intereses significativos
5. Fomentar el aprecio de la música, el arte, la literatura y otras experiencias estéticas

6. El desarrollo de la sensibilidad social
7. El desarrollo de un mejor ajuste personal y social
8. La adquisición de información importante
9. El desarrollo de la salud física
10. El desarrollo de una filosofía consecuente

Otra magnífica fuente de ideas sobre objetivos educativos generales es la *Taxonomía de objetivos educativos* (Bloom, 1956, y Krathwohl, 1964). El propósito original de la taxonomía fue abarcar las tres esferas conductuales: cognoscitiva, afectiva y psicomotora. El área cognoscitiva cubre los objetivos relacionados con el recuerdo o conocimiento de conocimientos y la solución de problemas. El dominio afectivo se refiere a cambios en los intereses, valores y actitudes y al desarrollo de la apreciación y el ajuste. El área psicomotora abarca los objetivos que se relacionan con las destrezas manuales y motoras. El trabajo de Bloom (1956) cubrió el dominio cognoscitivo, mientras que Krathwohl (1964) se ocupó del área afectiva. El trabajo sobre el dominio psicomotor todavía no ha sido publicado. Primero revisaremos muy brevemente el trabajo de Bloom y sus colaboradores en relación con nuestro interés por los objetivos. El grupo de Bloom se refirió en la siguiente forma al valor que sus resultados tienen para el maestro (Bloom, 1956):

La taxonomía puede ayudar a percibir el hincapié que se hace en ciertas conductas mediante un conjunto particular de planes educativos. Por tanto, un profesor, al clasificar las metas de una unidad de enseñanza, podrá dar cuenta de que todas caen dentro de la categoría taxonómica de la evocación de conocimientos. El examen de las categorías taxonómicas puede sugerirle que incluya algunas metas que se refieran a la aplicación de este conocimiento y al análisis de la situación donde se usa (página 2).

La clasificación de los objetivos educativos cognoscitivos consta de seis categorías principales:

- 1.00 Conocimiento
- 2.00 Comprensión
- 3.00 Aplicación
- 4.00 Análisis
- 5.00 Síntesis
- 6.00 Evaluación

Estas categorías principales constan de subcategorías, por ejemplo, entre las que pertenecen al conocimiento están las siguientes:

- 1.00 Conocimiento
 - 1.10 Conocimiento de específicos
 - 1.11 Conocimiento de la terminología
 - 1.12 Conocimiento de hechos específicos

- 1.20 Conocimiento de formas y procedimientos para manejar hechos concretos
- 1.30 Conocimiento de los universales y las abstracciones pertenecientes a un campo determinado

Abajo de cada encabezado aparece la definición y el análisis de su significado. Para que éste resulte más claro se incluyen ejemplos de metas educativas correspondientes a la categoría específica. Además, al final de cada sección se hace una revisión de los tipos de reactivos que pueden usarse para examinar el logro de las diferentes facetas de cada objetivo. Se recomienda ampliamente la *Taxonomía de objetivos educativos* de Bloom (1956), libro que es una excelente fuente para preparar unidades de estudio y tests.

Krathwohl (1964) y su comité prepararon las siguientes categorías principales para el dominio efectivo:

- 1.0 Recibir. Sensibilidad a la presencia de ciertos estímulos.
- 2.0 Responder. Atención activa a los estímulos, por ejemplo, responder a las reglas y las prácticas.
- 3.0 Valorar. Creencia consistente y actitud positiva hacia un fenómeno.
- 4.0 Organización. Organizar, interrelacionar y analizar diferentes valores que vayan al caso.
- 5.0 Caracterización conforme a un valor o un concepto de valor. La conducta es guiada por el valor.

Para un análisis detallado de estos objetivos, véase la *Taxonomía de objetivos educativos* de Krathwohl.

Objetivos específicos

En el área de objetivos específicos nos interesa el campo específico de estudio. Este abarca las destrezas, conceptos, hechos, principios, etcétera. Un buen método para establecer los objetivos educativos consiste en describir los elementos anteriores.

La definición de los objetivos no debe pecar de laboriosa. No se abruma con la tarea; realícela en provecho propio. Esto significa elaborar, en forma práctica, objetivos para unidades de estudio, incluyendo métodos de evaluación. Por ejemplo, consideremos una unidad de aritmética.

Objetivo general

- a) Promover la destreza en la adición y sustracción con decenas y centenas de dígitos.

Objetivos específicos

a) Adición

1. Problemas que requieren de llevar unidades.
2. Problemas que no requieren de llevar unidades.
3. Uso de puntos decimales.

b) Sustracción

1. Problemas que requieren de tomar unidades prestadas.
2. Problemas que no requieren de tomar unidades prestadas.
3. Uso de puntos decimales.

Traducamos nuestros "objetivos generales" y "específicos" de aritmética a preguntas de examen efectivas:

Adición		Sustracción	
1.	34	30	46
	+34	-19	-27
	67	46	634
	+26	-488	-5797
	834	634	7526
	+287		
	5468		
	+7679		
2.	65	48	84
	+32	-32	-63
	15	84	975
	+61	-960	-2214
	213	975	3345
	+381		
	8065		
	+1924		
3.	42	5.7	6.8
	+3.3	-4.5	-1.9
	5.33	6.8	8.43
	+6.11	-7.99	-565.2
	+1.33		
	88.44		
	+16.67		

La figura 41 contiene una copia de un test correspondiente a los objetivos educativos en materia de aritmética.

Objetivos	I. Las bases de nuestro Gobierno Constitucional (30 por ciento de los reactivos)	II. Principios y desarrollo de nuestro Gobierno Constitucional (70 por ciento de los reactivos)
1. Conocer los términos y el vocabulario (10%)	1. Derechos inalienables 2. Tiranía 3. Compromiso 4. Confederación (1 o 2 reactivos)	Préstbulo Legislativo Ejecutivo Judicial Federalismo Gobiernos unitarios Cláusula elástica Poderes residuales Soberanía Poderes concurrentes Sistema de balance y revisión Sufragio (4 reactivos)
2. Conocer fechas, acontecimientos, personas y lugares (5%)	1. Artículos de la Confederación—1781 2. Carta Magna—1215 3. Carta de Derechos—1689 5. Declaración de la Independencia—1775 4. Petición de Derechos—1628 6. Establecimiento de la Convención Constitucional —miembros, acuerdos 7. Ley de 1787 (0-1 reactivo)	Ratificación de la Constitución—1788 Papeles federales. Estado y propósitos (2 reactivos)

Objetivos	I. Las bases de nuestro Gobierno Constitucional (30 por ciento de los reactivos)	II. Principios y desarrollo de nuestro Gobierno Constitucional (70 por ciento de los reactivos)
3. Conocer generalizaciones, conceptos y principios (35%)	1. Disposiciones de la Carta Magna, Carta de Derechos, Petición de Derechos 2. Principios incluidos en la Declaración de Independencia 3. Debilidades de los artículos de la Confederación 4. Disposiciones de la ley de 1707 (6 o 7 reactivos)	1. Restricciones sobre los estados y el gobierno nacional 2. Poderes que pertenecen solamente al estado 3. Poderes que pertenecen solamente al gobierno nacional 4. Poderes que comparten los estados y el gobierno nacional 5. Disposiciones de los primeros VII artículos 6. Cambios y reformas de la Constitución: cómo se hicieron 7. Disposiciones de la Carta de Derechos (14 o 15 reactivos)
4. Señalar el desarrollo de nuestra forma nacional de gobierno (20%)	1. Identifique las influencias y tendencias que contribuyeron al establecimiento de juntas representativas en las colonias 2. Explicación del desarrollo del gobierno representativo en las colonias entre 1607 y 1776 3. Explique la manera como la ley de 1787 afectó al desarrollo de nuestro país (3 o 4 reactivos)	1. Identifique los sucesos, fuerzas y temores que condujeron a la separación de los poderes del gobierno federal 2. Identifique las fuerzas que condujeron a las reformas de la Constitución (8 o 9 reactivos)
5. Puede expresar generalizaciones y conceptos con sus propias palabras (10%)	1. Igualdad del hombre 2. Los gobiernos establecen sus justos poderes con el consentimiento de los gobernados 3. Explique por qué los artículos de la Confederación no resultaron un completo fracaso (1 o 2 reactivos)	1. Explique el significado de las reformas constitucionales 2. Nuestro gobierno como república democrática (4 reactivos)
6. Puede señalar relaciones, similitudes y diferencias (10%)	1. Relación entre los orígenes nacionales de los colonos y las disposiciones del Tratado de Mayflower 2. Abusos publicados en la Declaración de Independencia y si aún se cometen actualmente en nuestro país (1 o 2 reactivos)	1. Relación entre el preámbulo de la Constitución y los artículos de la Confederación 2. Similitudes y diferencias entre el gobierno nacional y el local 3. Similitudes y diferencias entre el Plan Virginia y el Plan de New Jersey y la Constitución cuando se adoptó (4 reactivos)
7. Puede aplicar generalizaciones y principios a situaciones nuevas (10%)	1. Identifique ventajas y desventajas de organizaciones de tipo confederado como las Naciones Unidas (1 o 2 reactivos)	1. Señale las diferencias esperadas entre nuestro gobierno y un gobierno organizado bajo otros principios 2. Identifique actos que violen la Carta de Derechos y el derecho violado por cada acto 3. Prediga el probable resultado si la Constitución no fuera flexible 4. Explique el proceso por el cual Puerto Rico pudo convertirse en estado 5. Especifique si las leyes propuestas están en conflicto con la Constitución 6. Identifique errores y procesos parlamentarios (4 reactivos)

Tiempo total para el test: 50 minutos
 Número total de reactivos del test: 60. (Asunto I-18 reactivos; Asunto II-42 reactivos.)

FIGURA 41. Reproducción del texto de una unidad sobre *Cómo funciona nuestro Gobierno*. (Tomado de Robert L. Thorndiko y Elizabeth Hagen, *Measurement and Evaluation in Psychology and Education*, © 1961, de John Wiley & Sons, Inc.) Reimpreso con permiso de John Wiley & Sons, Inc.

PASOS EN LA ELABORACIÓN Y APLICACIÓN DE TESTS

Hasta el momento hemos hablado de estos aspectos en términos generales. Ahora describiremos, paso a paso, el proceso real de construcción y aplicación de tests.

1. Etapa preparatoria.

- a) Reúna todas las materiales de enseñanza que tengan relación con el test proyectado, por ejemplo, libros, notas y descripción del contenido u objetivos de las unidades.
- b) Enumere los objetivos que desea medir.

2. Etapa de elaboración.

- a) Planee el tipo de test que va a usar y su formato general.
 - b) Haga un borrador los reactivos del test.
 - c) Fije la extensión del test. ¿Qué tan grande desea que sea el test: de todo el curso o de una parte?
 - d) Revise los reactivos. Hágalo días después de haber redactado el borrador. Elimine los reactivos que no vengán al caso y corrija los demás.
 - e) Ordene los reactivos. Estime la dificultad de los reactivos y coloque el más fácil al principio para no desanimar a los estudiantes y proporcione a los malos estudiantes la motivación para que sigan haciendo el test.
 - f) Instrucciones. Decida lo que desea que hagan los estudiantes y comuníquelos.
 - g) Precise el procedimiento de calificación. Asegúrese haber escrito las respuestas correctas antes del examen. Hágalo tanto para los tests de ensayo como para los reactivos objetivos.
3. Etapa de aplicación.
 - a) Tiempo. Cuquiera que sea el tiempo estipulado, sea consistente. No diga, por ejemplo: "Ustedes tendrán treinta minutos para contestar el test" y después conceda una hora a los estudiantes más lentos. No es equitativo para los que trataron de completar la tarea en el tiempo asignado. Normalmente, conceda suficiente tiempo para que por lo menos termine el test: el 90 por ciento de los examinados.
 - b) Condiciones físicas. Trate de mantener condiciones de tranquilidad, sin perturbaciones exteriores. No se preocupe por el salón de exarzan, no golpee con el lápiz, no converse con otros profesores, etcétera.

- c) Aplicación. No sea hostil en el tono o en la forma de las instrucciones. Su voz debe ser clara y las instrucciones deben ser comprendidas fácilmente por todos los estudiantes. Un examen es difícil para cualquiera.

EVALUACIÓN DE SU TEST

El lector ya está enterado, por nuestro estudio anterior, lo que es un buen instrumento de medida. En esta sección repetiremos algunos de los factores que se relacionan con los tests hechos por el profesor. (1965, págs. 281-307) ha resumido las características de un buen test señalando las siguientes diez cualidades:¹

1. *Perinencia.* ¿Su test mide sus objetivos educativos?
2. *Alicner.* ¿Qué tan cerca está su test del "ideal"? ¿Sus reactivos reflejan los objetivos expresados?
3. *Eficiencia.* ¿En cuánto tiempo se aplica su test? ¿Es demasiado? Debe lograrse una especie de término medio entre el tiempo disponible para el examen y la calificación y las otras necesidades.
4. *Objetividad.* Debe haber acuerdo entre los expertos sobre la respuesta correcta o la mejor respuesta a una pregunta. La objetividad no se refiere al tipo de test (por ejemplo, reactivos de opción múltiple o completamiento) sino a la forma de calificar.
5. *Especificidad.* ¿Está usted examinando el campo de estudio presentado en la clase? Es decir, el test discrimina entre los estudiantes que han aprendido ese campo de estudio y los que sólo pueden adivinar?
6. *Dificultad.* La dificultad de los reactivos debe estar al nivel del grupo que se va a examinar. Generalmente podemos considerar apropiado un test cuando cada uno de sus reactivos es resuelto por la mitad de los estudiantes.
7. *Discriminación.* Un reactivo de examen discrimina si los estudiantes correctamente más estudiantes buenos que malos.
8. *Confabilidad.* ¿Alide el test consistentemente lo que se pretende?
9. *Egüidad.* ¿Tiene cada estudiante la misma oportunidad de demostrar sus conocimientos? El test debe construirse de manera que cada estudiante pueda por igual demostrarlo.

¹ Se han hecho ligeras modificaciones en las descripciones, pero no en las cualidades, por ejemplo, en la *perinencia*.

10. *Rápidez.* ¿Están influidas las puntuaciones por la rapidez de la respuesta, y si es así, ¿hasta qué punto? En el test de aprovechamiento, la rapidez debe desempeñar un papel mínimo en la determinación de la puntuación de un estudiante. Debe haber bastante tiempo para que casi todos terminen el test.

TIPOS DE TESTS HECHOS POR EL PROFESOR

Existen básicamente dos tipos de tests hechos por el profesor. Son el test de ensayo y el objetivo. No obstante, hay muchas formas de tests objetivos mientras que los de ensayo se restringen generalmente al formato de "respuesta corta" o "análisis". La siguiente lista de tests objetivos hechos por el profesor se cita para mostrar la diversidad de formatos disponibles. Algunos de ellos se analizarán en los capítulos siguientes.

TESTS OBJETIVOS:

1. Recuerdo simple
 - a) Básico
 - b) Tipo de problema
 - c) Mapas, cartas y cosas por el estilo
2. Completamiento
 - a) Básico
 - b) Pareamiento
 - c) Analogías
 - d) Mapas, cartas y similares
3. Respuesta alternativa
 - a) Básico
 - b) (de) dos cláusulas
 - c) Tres alternativas
 - d) Inversa
 - e) Con corrección
 - f) Con calificación (o requisito)
 - g) Con diagramas
 - h) Con analogías
4. Selección múltiple
 - a) Básico
 - b) Recordación

- c) Principio común
 - d) Resultados
 - e) Causas
 - f) Cartas, mapas, etcétera.
 - g) Analogías
5. Apareamiento
- a) Básico
 - b) Tres columnas
 - c) Lista maestra
 - d) Analogías
6. Reordenación (o reajuste)
- a) Cronológico
 - b) Orden de importancia
 - c) Orden de dificultad
 - d) Longitud, peso, lógica, etcétera.

En los capítulos siguientes analizaremos algunos de los tipos anteriores de tests hechos por el profesor. El lector ya está familiarizado con algunos de ellos por el capítulo sobre tests de aprovechamiento.

TESTS DE ENSAYO EN CONTRASTE CON TESTS OBJETIVOS

No existen reglas definitivas para saber si el test de ensayo o el objetivo es mejor, ni tampoco para determinar cuándo es más ventajoso preferir uno de ellos. El profesor está en la mejor posición para seleccionar una forma en lugar de la otra. La selección se basa en los requisitos particulares y en las necesidades de la situación escolar y, más específicamente, en lo que el profesor desea medir y en el tiempo disponible para ello. La tabla 26 presenta un resumen de unas cuantas de las características principales de los dos tipos de tests.

El estudiante de medición debe recordar que ninguno de estos tests aislados, es completamente satisfactorio para medir el progreso académico. Cada uno tiene sus propias ventajas y desventajas y es responsable del maestro usar el test adecuado en las circunstancias apropiadas y para el propósito adecuado.

LOS PROFESORES, LOS TESTS Y LA REALIDAD

Se ha tratado en este capítulo de dar algunas orientaciones para elaborar tests. En los capítulos siguientes se presentarán ejemplos y experiencias precisas para elaborar tests de ensayo y objetivos. No obstante, debe advertirse que por razones de espacio, nuestro análisis será limitado.

TABLA 26. Comparación entre tests de ensayo y objetivos

	Ensayo	Objetivos
Capacidades medidas	Demanda que el estudiante se exprese con sus propias palabras usando informaciones extraídas de sus conocimientos y antecedentes.	Requiere que el estudiante seleccione las respuestas correctas de las opciones dadas, o que proporcione una respuesta limitada a una palabra o frase.
Alcanza	Cubre solamente un campo limitado de conocimientos en cualquier prueba. Las preguntas de ensayo necesitan de tanto tiempo para contestarse que relativamente pocas pueden contestarse en un período dado. También, el estudiante que es especialmente sobresaliente podrá evitar discutir puntos de los que no está muy seguro.	Un test cubre un amplio campo de conocimientos. Ya que las preguntas objetivas pueden contestarse rápidamente, un test puede contener muchas preguntas. Una gran extensión ayudará a proporcionar medidas confiables.
Incentivo para los alumnos	Anima a los estudiantes a aprender a organizar sus propias ideas y a expresarlas efectivamente.	Anima al estudiante a formarse una base amplia de conocimientos y capacidades.
Fácil de preparar	Solamente se requiere escribir unas cuantas preguntas para el test. Las tareas deben estar claramente definidas, ser lo suficientemente generales como para ofrecer libertad, y suficientemente específicas para marcar límites.	Exige muchas preguntas para un test. Deben evitarse las ambigüedades y las indicaciones de la respuesta correcta en la redacción. Los distractores deben incluir las nociones erróneas.

TABLA 26. (Continúa)

	Ensayo	Objetivos
Cualificación	Generalmente se requiere de mucho tiempo para calificar.	Pueden calificarse rápidamente.
	Permite que los profesores comenten directamente el proceso de razonamiento de los alumnos individuales. Sin embargo, una respuesta puede ser calificada en forma diferente por distintos profesores o el mismo profesor en diferentes ocasiones.	Las respuestas se califican generalmente sólo como correctas o incorrectas, pero la calificación es muy exacta y consistente.

Tomado de *Making the classroom tests: A Guide for Teachers*, Educational Testing Service Evaluation and Advisory Service Series núm. 4. © Primera edición, Copyright 1959 de Educational Testing Service. Segunda edición: 1961. Reproducción autorizada por Educational Testing Service.

Recuérdese también que ningún test aislado diseñado por el profesor debe ser la única base para tomar decisiones educativas importantes. El desempeño diario en la clase, las puntuaciones en tests estandarizados y el aprovechamiento académico medido por una serie de tests, dados por el profesor, todo esto deberá tomarse en cuenta en el estudio completo de la evaluación del estudiante. Como señala Nunnally: "llegar a conclusiones importantes sobre los estudiantes fundándose en un solo test diseñado por el profesor sería tan insensato como que el cazador abandonara su tarea porque en la primera palada no estuvo oro" (página 106).

REFERENCIAS

Allin, W. M. *The Story of the eight-year study, with conclusions and recommendations*. Nueva York: Harper & Row, 1942.
 Bloom, B. S. (prep. ed.) *Taxonomy of educational objectives. Handbook I: The cognitive domain*. Nueva York: David McKay, 1956.
 Ebel, R. L. *Measuring educational achievement*. Englewood Cliffs, N. J.: Prentice-Hall, 1965.
 Kariswohl, D. R., Bloom, B. S. y Masia, B. D. *Taxonomy of educational objectives. Handbook 2: The affective domain*. Nueva York: David McKay, 1954.
 Linn, A. J. *Measurement and evaluation of learning: A handbook for teachers*. Dubuque, Iowa: Wm. C. Brown, 1957.

Noll, V. H. *Introduction to educational measurement*. (2^a ed.) Boston: Houghton Mifflin, 1965.

Nunnally, J. C. *Educational measurement and evaluation*. Nueva York: McGraw-Hill, 1964.

United States Department of Interior, Bureau of Education. *Cardinal principles of education*. Boletín núm. 38, 1918.

INSTRUMENTO PARA LA OBSERVACIÓN SISTEMÁTICA DEL NIVEL DE DESARROLLO DEL ALUMNO EN EL 1.º EGB

(Anexo 4. Instrumento de observación sistemática)

PROCEDENCIA	MISES DE NACIMIENTO												
	EDAD DEL ALUMNO CUANDO EMPLEA EL CURSO												
ESCOLARIDAD	En. 0.8	F. 0.7	Mz. 0.6	Ab. 0.5	Ms. 0.4	Ja. 0.3	Jl. 0.2	A. 0.1	Sr. 0	O. S.11	N. S.10	D. S.0	
No escolarizados													
En guarderías													
En centros Preescolares Autónomos	desde P-2												
	desde P-1												
En centros Preescolares Integrados	desde P-2												
	desde P-1.												

- El área superior izquierda indica: mayor inadaptación al ambiente escolar. Coincidiendo: mayor edad y menos escolarización.
- El área inferior izquierda indica: mayor adaptación al ambiente escolar. Coincidiendo: mayor edad y más escolarización.
- El área superior derecha indica: menor apoyo a la maduración personal. Coincidiendo: menor edad y menor escolarización.
- El área inferior derecha indica: mayor apoyo a la maduración personal. Coincidiendo: menor edad y mayor escolarización.

GALINDO CHAVANEL, Ma. Agustina. "Instrumento para una observación sistemática", en: Cómo coordinar el nivel de preescolar y el ciclo inicial. Madrid, Editorial Escuela Española, 1984. pp. 165-177.

OBSERV	APOYO NECESARIO	
	Observación (1)	Programa de compensación (2)
— Conocimiento de:		
— Agudeza y utilización:		
— Soltura en la manipulación:		
— Pensamiento infantil:		
— Resuelve situaciones:		
— Imagina:		
— Uso del lenguaje:		
— Pronunciación:		
— Construcción:		
— Comprensión:		
— Conversación:		
— Expresión:		
— Corporal:		
— Dinámica o plástica:		
— Musical:		
— Plástica:		
— Fantástica:		

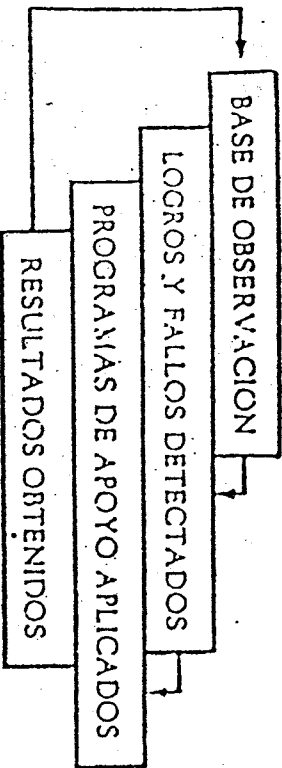
(1) Expresado en código
(2) Expresado en código

OBSERV	COMUNICACIONES NECESARIAS	
	Observación (1)	Programa de compensación (2)
— Hábitos adaptados:		
— Higiéncicos:		
— control:		
— Aseo:		
— Adaptación:		
— Lo com:		
— Participa:		
— Control:		
— Trabajo en grupo:		
— Se mueve:		
— Intenta:		
— Actitudes:		
— Individual:		
— Se mueve:		
— Es resis:		
— Se inicia:		
— Se mueve:		
— Sociales:		
— Ignora:		
— Partici:		
— Asimilación:		
— Es sincero:		
— Renuncia:		
— Comparte:		

(1) Expresado en código
(2) Expresado en código

INSTRUMENTO PARA UNA OBSERVACION SISTEMATICA

- La observación sistemática, en manos del maestro, supone, prepara y completa un diagnóstico.
- Se acompaña un inventario de pautas observacionales, con un documento de apoyo al equipo de profesores de párvulos y ciclo inicial, cuando se proponen llevar a cabo la observación sistemática de sus alumnos.
- De las observaciones hechas pueden quedar constancia en una ficha individual abierta, y servir de base para planear el trabajo del aula.



¿COMO OBSERVAR?

1.º UTILIZAR EL INSTRUMENTO

- ¿Cómo determinar si la conducta observada en un niño es normal?
- El equipo de profesores debe plantearse estas preguntas y tomar acuerdos antes de empezar la tarea de observación.
- En todo caso, tendrán que tener en cuenta:
 1. La edad de los alumnos, ya que en cada edad:
 - Aparecen nuevos comportamientos.
 - Desaparecen ciertos modos de conducta infantiles, propios de fases anteriores.
 2. El comportamiento individual y social del grupo de niños de la misma edad:
 - Comparar cada caso y observar si su conducta es reiterativa, o si hay ausencias de otras conductas que son respuesta propia a una situación.
 - Cuando un niño adopta con demasiada frecuencia un modo de respuesta ante variadas situaciones, observar si presenta ausencias de conductas que son propias de su edad y medio en el que se desenvuelve.

2.º MANEJO DE LA FICHA INDIVIDUAL Y ABURRIDA PARA CADA ALUMNO

1. Anotar en fichas reunidas los rasgos más significativos que se observan en cada aspecto respecto del niño observado.
2. Señalar los logros y fallos. Conductas que más llamaron la atención por su frecuencia-ausencia o características particulares del grupo.
3. Incluir el programa de apoyo aplicado y los resultados obtenidos.
4. Recoger este material en carpeta de fácil manejo.
5. Incluir en todos los aspectos una interpretación del hecho.

3.º INVENTARIO DE PAUTAS OBSERVACIONALES

«A través del proceso de observación se hubrán relacionado algunas conductas del niño a las que antes no encontrabas explicación. Se hubrán puesto de manifiesto sus aptitudes específicas para realizar determinadas tareas y que hemos de seguir potenciándolas, así como también se habrán manifestado sus retrasos, sus problemas, y que hemos de afrontarlos a través de actividades adecuadas a sus de subanarías y conseguir, en la medida de lo posible, un desarrollo normal. De esa forma, los fracasos escolares disminuirán notablemente. Merece la pena el esfuerzo que supone todo el proceso de la observación sistemática.» (Esmeralda Rubio: Documento ciclarislado. Curso de Especialización del Profesorado de Educación Preescolar. Tarragona, Enero de 1980.)

PAUTAS OBSERVACIONALES DEL ASPECTO FISICO-MOTOR

A) APARIENCIA

- Peso y estatura proporcional: aspecto armonioso.
- Piel sin manchas, heridas, granos, etc.
- Su aspecto es saludable.
- Defectos físicos.
- Defectos sensoriales: visuales (ambliopía), auditivos (hipoacusia).
- Deficiencias físicas: sus actividades revelan un bajo nivel energético, se cansa con mucha facilidad.
- Hábitos higiénicos con respecto al cuerpo, a la ropa, a la alimentación.
- Falta frecuentemente a clase por enfermedad.

B) COORDINACIÓN MOTORA

I. LOCOMOCION

- Es capaz de caminar, correr, saltar.
- Circula por el aula con facilidad, con destreza.
- Tropieza continuamente.
- Se cae sin razón.
- Choca con todo.
- Es capaz de saltar con ambos pies, con un solo pie.
- Mantiene bien el equilibrio.

II. II. POSTURA

- Posición correcta cuando está sentado.
- Posición correcta al caminar.
- Altura desigual de los hombros al caminar.
- Recuesta la cabeza en un brazo o mano cuando trabaja.
- Se sienta mal, «tirador».
- Se sienta encogido.

III. EXPRESION DINAMICA

- Sus movimientos son coordinados y precisos.
- Cómo es su sentido del ritmo.
- Sus gestos y movimientos son expresivos.
- Es capaz de dramatizar estados de ánimo.
- Es capaz de escenificar cuentos.

IV. MANIPULACION

- a) **Uso de las manos**
 - Es zúrdo, diestro, ambidiestro.
 - Vuelve las páginas de los libros sin romperlas, sin arrugarlas o mancharlas.
 - Demuestra habilidad para recortar, pegar, modelar, rasgar, plegar, atornillar, etc.
 - Es capaz de movimientos que exigen una coordinación exacta, como clavar un clavo, atar los cordones de los zapatos, etc.
 - Sabe abrocharse los botones de la bata.
- b) **Uso del lápiz**
 - Prehensión palmar, con el lápiz vertical.
 - Coge el lápiz correctamente entre los dedos índice y pulgar.
 - Dibuja y escribe con suavidad, haciendo deslizar suavemente el lápiz.

- Presiona mucho, tanto que, a veces, rompe la mina del lápiz o rasga el papel.
- Le gusta mucho dibujar y pintar.
- Al colorear se sale del contorno de las figuras.

V. HABITOS NERVIOSOS

- Se mete los dedos en la nariz.
- Lleva a la boca cualquier objeto.
- Se chupa los dedos.
- Se roe las uñas.
- Se muerde los labios.
- Se arranca los cabellos.
- Se golpea o lastima el mismo.
- Contrae desordenadamente la cabeza.
- Contrae los brazos, las manos o los pies.
- Se queda llamativamente quieto mucho tiempo.
- Hace guiños.
- Muerde el lápiz.
- Balbucea una pierna.
- Se rasca frecuentemente.

PAUTAS OBSERVACIONALES DEL ASPECTO EMOCIONAL

I. CARACTERISTICAS DE SU TONALIDAD AFECTIVA

- Cómo fueron sus reacciones emocionales en los primeros días de su asistencia al preulario.
- En sus reacciones predomina la alegría, la agresión, el enojo, la inquietud, la apatía, etc.
- Es tímido y retraído.
- Es muy sensible a las reacciones emocionales. Se confunde con facilidad.
- Su comportamiento es inestable, descontrolado.

II. MANIFESTACIONES AFECTIVAS

- Es dado a frecuentes demostraciones: besos, abrazos, caricias, etc.
- Parece ávido de afecto, le gusta que le acaricien.
- Vive colgado del cuello de los demás.
- Da y solicita frecuentes pruebas de afecto.
- Sólo rara vez solicita demostraciones de afecto.
- Nunca acaricia y evita estas demostraciones.
- Cuando se siente lastimado le gusta que le hagan caricias.

III. REACCIONES EMOCIONALES A LA FRUSTRACION

- Es sensible a la censura: llora, se pone agresivo, etc.

- o Es rechazado por sus compañeros.
- o Se queja constantemente, y siempre está reusando compañeros.
- o Sus relaciones con otros miembros de la clase se limitan por los contactos físicos primarios e intercambio de objetos.
- o Sus relaciones se caracterizan por: contactos, preguntas, conversaciones, etc.
- o Ayuda espontáneamente a sus compañeros a los más pequeños.
- o Es afable y cordial con personas con quienes vuelve reservado cuando se enfrentan con alguien que no conoce.
- o Es capaz de ser el jefe de un juego o ser el mismo.
- o Es capaz de iniciativas, sabe colaborar.
- o Se ofrece voluntariamente para realizar trabajos.

III. RELACIONES SOCIALES CON EL MAESTRO

- o Es muy amable con la profesora, quiere ser como ella.
- o Conoce y respeta la autoridad de la profesora.
- o Obedece a la maestra con naturalidad.
- o Es capaz de escuchar al profesor, de mantenerse con él.
- o Explica a la señorita experiencias familiares y escolares.

PAUTAS OBSERVACIONALES DEL ASPECTO INTELECTUAL

I. RESPECTO AL TRABAJO ESCOLAR

- o Se cansa con facilidad.
- o Es incapaz de concentrarse en su trabajo.
- o Participa activamente en el trabajo escolar.
- o Muestra interés por las tareas que se realizan.
- o Es muy rápido en su trabajo.
- o Su atención es buena.
- o Tiene buena memoria.
- o En su modo de trabajar es organizado.
- o Mantiene su material pronto para ser utilizado.
- o Es cuidadoso con respecto a su materia.
- o Cuida su presentación de sus tareas.

II. OBSERVACION DE LOS PROCESOS

- o Tiene dificultades en la ejecución de actividades reales, como plegar, cortar, rasgar.
- o Dificultades en los ejercicios sensoriales como: oír, tocar, sentir, etc.

- o Ante una frustración se pone muy agresivo.
- o Su agresividad es física: empujones, pellizcos, mordidas, golpes.
- o Su agresividad es verbal: insulta, grita, amonesta, etc.
- o A la frustración reacciona con autoagresividad.
- o Ante las desilusiones se inhibe, llora, etc.

IV. SENSIBILIDAD ARTISTICA

I. EXPRESION PLASTICA

- o Sus dibujos se distinguen por alguna característica especial.
- o Sus dibujos son estereotipados.
- o Es muy creador en sus obras.
- o Utiliza todos los colores o sólo algunos.
- o Siente placer en pintar y dibujar.
- o Le gusta trabajar con plastilina o barro.

II. EXPRESION MUSICAL

- o Tiene buen sentido del ritmo.
- o Tiene agudeza y precisión auditiva.
- o Sabe valorar el silencio para aprender a escuchar.
- o Conoce canciones sencillas y juegos.

PAUTAS DE OBSERVACION SISTEMATICA DEL ASPECTO SOCIAL

I. AJUSTE A LAS NORMAS ESCOLARES (SITUACION)

- o Llega puntualmente al parvulario.
- o Deja su abrigo en el lugar correspondiente.
- o Recoge su bata y sabe ponérsela y abotonarla.
- o Es disciplinado.
- o Es cuidadoso, tanto con el propio material como con el material de sus compañeros.
- o Una vez utilizado, es capaz de recogerlo, ordenarlo y colocarlo en su sitio.
- o Obedece las órdenes y respeta las reglas establecidas para ir al lavabo, para comer la merienda, etc.
- o Terminada la clase, deja su silla y mesa bien colocadas, y el lugar donde está sentado, limpio de papeles.
- o Utiliza la papelería.

II. RELACIONES SOCIALES CON LOS COMPAÑEROS

- o Parece gustar de los contactos sociales, procurando siempre hacerse amigos.
- o Comparte sus cosas con los demás.
- o Tiene muchos pocos, ningún amigo.
- o Evita a los compañeros, aislándose.

- Dificultades en los ejercicios sensoriales del ojo: conocimiento de formas, colores, matices, gamas.
- Distingue la simetría de los objetos.
- Conoce las diversas posiciones de los objetos relacionados entre sí.
- Tiene problemas de lateralización y estructuración del espacio con respecto al esquema corporal.

III. DESARROLLO DEL LENGUAJE

- Comprende órdenes cada vez de complejidad creciente.
- Sabe escuchar.
- Pronuncia las palabras adecuadamente.
- Utiliza los vocablos de forma adecuada.
- Tiene dificultades en algún fonema en particular.
- Su vocabulario activo y pasivo es amplio.
- Estructura bien las frases.
- Sabe explicar alguna experiencia familiar o personal.
- Sabe dialogar.
- Es capaz de interpretar imágenes y escenas sencillas.
- Sabe recitar pequeñas poesías.
- Le gustan los trabalenguas.
- Simula hablar por teléfono.
- Hace preguntas oportunas e inteligentes.
- Sabe interpretar los libros de imágenes.

IV. DESARROLLO LOGICO-MATEMATICO

- Clasifica objetos según un criterio definido de antemano: color, forma, tamaño, grosor, etc.
- Es capaz de resolver problemas de la vida real.
- Ordena objetos por tamaños.
- Conoce los distintos términos que indican posición: arriba-abajo...
- Sabe puntuar el borde de las figuras.
- Observar si une bien sentido y grafismo de los números.
- Observar si se dan confusiones de números de sonidos semejantes.
- Observar si se dan confusiones de números de forma o grafismo semejante.
- Confusión de números simétricos.
- Confusión de los signos sumar, restar, igual.
- Distingue correctamente la escritura de: 2 y S, 3 y E.
- Sabe colocar correctamente cuando se le dice que se ponga el tercero o el quinto.
- Sabe resolver sencillos problemas, del estilo de: si tienes ocho caramelos y te comes tres, ¿cuántos te quedan?
- Comete errores al escribir las cifras del 1 al 9 en orden creciente. ¿en qué lugar los comete?
- Obedece órdenes en las que se le indica que traiga un número determinado de cosas.

CAMPOS DE OBSERVACION DE CONDUCTA

- Amplitud del tiempo de observación
- Edades en las que se observa
- Profesores/es

FALTAS OBSERVACIONES	
Las Actos Personales (1)	<ul style="list-style-type: none"> • Apariencia externa • Locomoción • Presura • Expresión dinámica • Manipulación • Uso de las manos • Uso del lápiz • Hábitos nerviosos • Aspecto emocional • Hábitos posturales • Hábitos de atención • Hábitos de conducta • Hábitos de higiene • Hábitos de alimentación • Hábitos de sueño • Hábitos de vestimenta • Hábitos de salud
Faltas directas (1)	
Programa de apoyo (1)	
Resultados (1)	

NIVEL DE ACQUISICIONES, DESEMPEÑOS Y DESEMPEÑOS

Referencia	Paralelo 1	Paralelo 2	1º EGB	2º EGB
ACTITUDES: — Individuales — Sociales				
HÁBITOS: — Higiénicos — De trabajo				
DESBREZAS: — Educación dinámica — Expresión artística				
ASIMILACION DE VALORES: — Morales — Religiosos				
CONOCIMIENTOS: Lenguaje: — Comprensión oral — Conversación				

— Lectura: — Expresión y entonación — Comprensión				
— Escritura: — Legibilidad — Nivel ortográfico — Composición				
Matemáticas: — Razonamiento — Cálculo operatorio — Resolución y planteo de problemas — Cálculo mental				
Entorno social y natural: — Comprensión — Retención				
Términos valorativos: — Inhibido, participativo, motivador — Descuidado, limpio, cuidadoso — Ausente, atento, diligente — Torpe, ágil, preciso — Inexpresivo, hábil, creativo — Suspicioso, sincero, abierto — Escaso, satisfactorio, avanzado				

Recolección de los datos

PROCESO DE INVESTIGACIÓN

Octavo paso

RECOLECTAR LOS DATOS

- Definir la forma idónea de recolectar los datos de acuerdo al contexto de la investigación.
- Elaborar el instrumento de medición.
- Aplicar el instrumento de medición.
- Obtener los datos.
- Codificar los datos.
- Archivar los datos y prepararlos para el análisis.

OBJETIVOS

Que el alumno:

- 1) Comprenda el significado de "medir" en ciencias sociales.
- 2) Comprenda los requisitos que toda medición debe cumplir: confiabilidad y validez.
- 3) Conozca los métodos para determinar la confiabilidad y validez de un instrumento de medición.
- 4) Comprenda los niveles de medición en que pueden ubicarse las variables.
- 5) Conozca los principales instrumentos de medición disponibles en ciencias sociales.
- 6) Esté capacitado para elaborar y aplicar diferentes instrumentos de medición.
- 7) Se encuentre habilitado en la preparación de datos para su análisis.

HERNANDEZ, Roberto, Carlos Fernández y Pilar Bautista. "Recolección de datos" en: Metodología de la investigación. Mc.Graw-Hill. pp.242-348.

Recolección de los datos

PROCESO DE INVESTIGACIÓN

Octavo paso

RECOLECTAR LOS DATOS

- Definir la forma idónea de recolectar los datos de acuerdo al contexto de la investigación.
- Elaborar el instrumento de medición.
- Aplicar el instrumento de medición.
- Obtener los datos.
- Codificar los datos.
- Archivar los datos y prepararlos para el análisis.

OBJETIVOS

Que el alumno:

- 1) Comprenda el significado de "medir" en ciencias sociales.
- 2) Comprenda los requisitos que toda medición debe cumplir: confiabilidad y validez.
- 3) Conozca los métodos para determinar la confiabilidad y validez de un instrumento de medición.
- 4) Comprenda los niveles de medición en que pueden ubicarse las variables.
- 5) Conozca los principales instrumentos de medición disponibles en ciencias sociales.
- 6) Esté capacitado para elaborar y aplicar diferentes instrumentos de medición.
- 7) Se encuentre habilitado en la preparación de datos para su análisis.

HERNANDEZ, Roberto, Carlos Fernández y Pilar Bautista. "Recolección de datos" en: Metodología de la investigación. Mc.Graw-Hill. pp. 242-348.

Este razonamiento nos hace sugerir que es más adecuado definir la medición como "el proceso de vincular conceptos abstractos con indicadores empíricos", proceso que se realiza mediante un plan explícito y organizado para clasificar (y frecuentemente cuantificar) los datos disponibles —los indicadores— en términos del concepto que el investigador tiene en mente (Caminos y Zeller, 1979, p. 10). Y en este proceso, el instrumento de medición o de recolección de los datos juega un papel central. Sin él no hay observaciones clasificadas.

La definición sugerida incluye dos consideraciones: La primera es desde el punto de vista empírico y se resume en que el centro de atención es la respuesta observable (sea una alternativa de respuesta marcada en un cuestionario, una conducta grabada vía observación o una respuesta dada a un entrevistador). La segunda es desde una perspectiva teórica y se refiere a que el interés se sitúa en el concepto subyacente no observable que es representado por la respuesta (Caminos y Zeller, 1979). Así, los registros del instrumento de medición representan valores observables de conceptos abstractos. Un instrumento de medición adecuada es aquel que registra datos observables que representan verdaderamente a los conceptos o variables que el investigador tiene en mente.

En toda investigación aplicamos un instrumento para medir las variables contenidas en las hipótesis (y cuando no hay hipótesis, simplemente para medir las variables de interés). Esa medición es efectiva cuando el instrumento de recolección de los datos realmente representa a las variables que tenemos en mente. Si no es así nuestra medición es deficiente y por lo tanto la investigación no es digna de tomarse en cuenta. Desde luego, no hay medición perfecta, es prácticamente imposible que representemos fielmente variables tales como la inteligencia, la motivación, el nivel socioeconómico, el liderazgo democrático, la actitud hacia el sexo y otras más; pero si debemos de acercarnos lo más posible a la representación fiel de las variables a observar, mediante el instrumento de medición que desarrollemos.

3.3. ¿QUÉ REQUISITOS DEBE CUBRIR UN INSTRUMENTO DE MEDICIÓN?

Toda medición o instrumento de recolección de los datos debe reunir dos requisitos esenciales: confiabilidad y validez. La confiabilidad de un instrumento de medición se refiere al grado en que su aplicación repetida al mismo sujeto u objeto, produce iguales resultados. Por ejemplo, si yo midiera en este momento la temperatura ambiental mediante un termómetro y me indicara que hay 22°C. Un minuto más tarde consultara otra vez y el termómetro me indicara que hay 5°C. Tres minutos después observara el termómetro y ahora me indicara que hay 40°C. Este termómetro no sería confiable (su aplicación repetida produce resultados distintos). Igualmente, si una prueba de inteligencia la aplico hoy a un grupo de personas y me proporciona ciertos valores de inteligencia; la aplico un mes después y me proporciona valores diferentes, al igual que en subsecuentes mediciones. Esa prueba no es confiable (analícense los valores de la figura 9.1, suponiendo que los coeficientes de inteligencia puedan oscilar entre 95 y 150). Los resultados no son consistentes; no se puede "confiar" en ellos.

FIGURA 11

EJEMPLO DE RESULTADOS PROPORCIONADOS POR UN INSTRUMENTO DE MEDICIÓN DE CONFIABILIDAD

	PRIMERA APLICACIÓN		SEGUNDA APLICACIÓN		TERCERA APLICACIÓN
Martha	130	Laura	131	Luis	140
Laura	125	Luis	130	Teresa	129
Arturo	118	Marco	127	Martha	124
Luis	112	Arturo	120	Rosa María	120
Marco	110	Chester	118	Laura	104
Rosa María	110	Teresa	118	Chester	108
Chester	108	Martha	115	Arturo	103
Teresa	107	Rosa María	107	Marco	101

La confiabilidad de un instrumento de medición se determina mediante diversas técnicas, las cuales se comentarán brevemente después de revisar el concepto de validez.

La validez, en términos generales, se refiere al grado en que un instrumento realmente mide la variable que pretende medir. Por ejemplo, un instrumento para medir la inteligencia válido debe medir la inteligencia y no la memoria. Una prueba sobre conocimientos de Historia debe medir esto y no conocimientos de literatura histórica. Aparentemente es sencillo lograr la validez. Después de todo —como dijo un estudiante— “pensamos en la variable y vemos cómo hacer preguntas sobre esa variable”. Esto sería factible en unos cuantos casos (como lo sería el “sexo” de una persona). Sin embargo, la situación no es tan simple cuando se trata de variables como la motivación, la calidad de servicio a los clientes, la actitud hacia un candidato político y menos aún con sentimientos y emociones, así como diversas variables con las que trabajamos en ciencias sociales. La validez es una cuestión más compleja que debe alcanzarse en todo instrumento de medición que se aplica. Kerlinger (1979, p. 138) plantea la siguiente pregunta respecto a la validez: ¿Está usted midiendo lo que usted cree que está midiendo? Si es así, su medida es válida; si no, no lo es.

La validez es un concepto del cual pueden tenerse diferentes tipos de evidencia (Wiersma, 1986; Gronlund, 1985): 1) evidencia relacionada con el contenido, 2) evidencia relacionada con el criterio y 3) evidencia relacionada con el constructo. Hablaremos de cada una de ellas.

1) Evidencia relacionada con el contenido

La validez de contenido se refiere al grado en que un instrumento refleja un dominio específico de contenido de lo que se mide. Es el grado en que la medición representa al concepto medido (Bohmstedt, 1976). Por ejemplo, una prueba de operaciones aritméticas no tendrá validez de contenido si incluye sólo problemas de resta y excluye problemas de

suma, multiplicación o división (Cammines y Zeller, 1979). O bien, una prueba de conocimientos sobre las canciones de "Los Beatles" no deberá basarse solamente en sus álbumes "Let it Be" y "Abbey Road", sino que debe incluir canciones de todos sus discos.

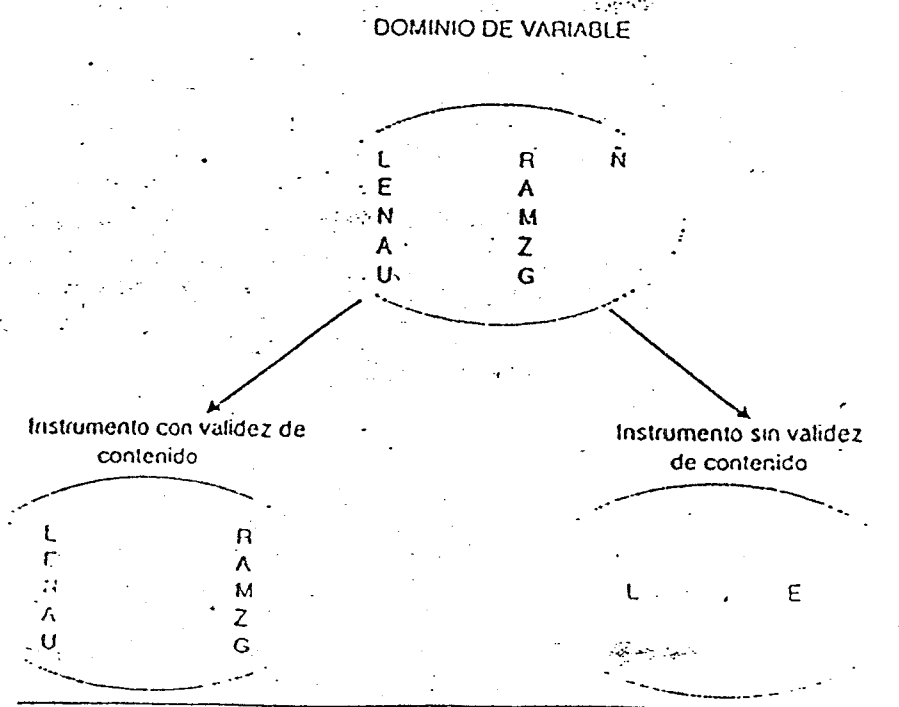
Un instrumento de medición debe *contener* representados a todos los ítems del dominio de contenido de las variables a medir. Este hecho se ilustra en la figura 9.2.

2) Evidencia relacionada con el criterio.

La *validez de criterio* establece la validez de un instrumento de medición comparándola con algún criterio externo. Este criterio es un estándar con el que se juzga la validez del instrumento (Wiersma, 1986). Entre los resultados del instrumento de medición se relacionen más al criterio, la validez del criterio será mayor. Por ejemplo, un investigador valida un examen sobre manejo de aviones, mostrando la exactitud con que el examen predice que tan bien un grupo de pilotos puede operar un acroplano.

FIGURA 9.2

ILUSTRACIÓN DE UN INSTRUMENTO DE MEDICIÓN CON VALIDEZ DE CONTENIDO VERSUS CON UNO QUE CARECE DE ÉSTA.



Si el criterio se refiere al *pasado*, se habla de *validad retrospectiva*. Los resultados del instrumento se comparan con los resultados en el mismo instrumento a punto de tiempo. Por ejemplo, un cuestionario para determinar las preferencias de voto por los distintos partidos contendientes, puede validarse aplicándolo tres o cuatro días antes de la elección y sus resultados compararlos con los resultados finales de la elección (si no hay fraude —desde luego—).

Si el criterio se fija en el *futuro*, se habla de *validad predictiva*. Por ejemplo, una prueba para determinar la capacidad administrativa de altos ejecutivos se puede validar comparando sus resultados con el futuro desempeño de los ejecutivos medidos.

3) Evidencia relacionada con el constructo

La *validad de constructo* es probablemente ³³ la más importante sobre todo desde una perspectiva científica y se refiere al grado en que una medición se relaciona consistentemente con otras mediciones de acuerdo con hipótesis derivadas teóricamente y que conciernen a los conceptos (o constructos) que están siendo medidos. Un *constructo* es una variable medida y que tiene lugar dentro de una teoría o esquema teórico.

Por ejemplo, supongamos que un investigador desea evaluar la *validad de constructo* de una medición particular, digamos una escala de motivación intrínseca: "el Cuestionario de Reacción a Tareas", versión mexicana (Hernández-Sampieri y Cortés, 1982). Estos autores sostienen que el nivel de motivación intrínseca hacia una tarea está relacionado positivamente con el grado de persistencia adicional en el desarrollo de la tarea (v.g., los empleados con mayor motivación intrínseca son los que suelen quedarse más tiempo adicional una vez que concluye su jornada). Consecuentemente, la predicción teórica es que a mayor motivación intrínseca, mayor persistencia adicional en la tarea. El investigador administra dicho cuestionario de motivación intrínseca a un grupo de trabajadores y también determina su persistencia adicional en el trabajo. Ambas mediciones son correlacionadas. Si la correlación es positiva y sustancial, se aporta evidencia para la *validad de constructo* del Cuestionario de Reacción a Tareas, versión mexicana (a la *validad* para medir la motivación intrínseca).

La *validad de constructo* incluye tres etapas:

- 1) Se establece y especifica la relación teórica entre los conceptos (sobre la base del marco teórico).
- 2) Se correlacionan ambos conceptos y se analiza cuidadosamente la correlación.
- 3) Se interpreta la evidencia empírica de acuerdo a qué tanto clarifica la *validad de constructo* de una medición en particular.

El proceso de *validación de un constructo* está vinculado con la teoría. No es posible llevar a cabo la validación de constructo, a menos que exista un marco teórico que soporte a la variable en relación con otras variables. Desde luego, no es necesaria

³³ La explicación se basa en Carnúmes y Zeller (1979).

una teoría sumamente desarrollada, pero si investigaciones que hayan demostrado que los conceptos están relacionados. Entre más elaborado y comprobado se encuentre el marco teórico que apoya la hipótesis, la validación de constructo puede arrojar mayor luz sobre la validez de un instrumento de medición. Y mayor confianza tenemos en la validez de constructo de una medición, cuando sus resultados se correlacionan significativamente con un mayor número de mediciones de variables que teóricamente y de acuerdo con estudios antecedentes están relacionadas. Esto se representa en la figura 9.3.

FIGURA 9.3

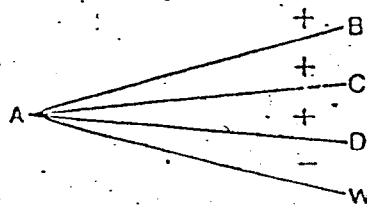
PRESENTACIÓN GRÁFICA DE UN INSTRUMENTO CON VALIDEZ DE CONSTRUCTO

Instrumento mide constructo "A"

Teoría

(Investigaciones hechas) encontraron que "A" se relaciona positivamente con "B", "C" y "D". Y negativamente con "W".

Si el instrumento mide realmente "A" sus resultados deben relacionarse positivamente con los resultados obtenidos en las mediciones de "B", "C" y "D"; y relacionarse negativamente con los resultados de "W".



El instrumento parece realmente medir "A".

Para analizar las posibles interpretaciones de evidencia negativa en la validez de constructo, se sugiere consultar a Cronbach y Mechl (1955) y Cronbach (1984).

VALIDEZ TOTAL = VALIDEZ DE CONTENIDO + VALIDEZ DE CRITERIO + VALIDEZ DE CONSTRUCTO

Así, la validez de un instrumento de medición se evalúa sobre la base de tres tipos de evidencia. Entre mayor evidencia de validez de contenido, validez de criterio y validez

de constructo tenga un instrumento de medición: este se acerca más a representar la variable o variables que pretende medir.

Cabe agregar que un instrumento de medición puede ser confiable pero no necesariamente válido (un aparato —por ejemplo— puede ser consistente en los resultados que produce, pero no medir lo que pretende). Por ello es requisito que el instrumento de medición demuestre ser confiable y válido. De no ser así, los resultados de la investigación no los podemos tomar en serio.

FACTORES QUE PUEDEN AFECTAR LA CONFIABILIDAD Y VALIDEZ

Hay diversos factores que pueden afectar la confiabilidad y la validez de los instrumentos de medición.

El primero de ellos es la improvisación. Algunas personas creen que elegir un instrumento de medición o desarrollar uno es algo que puede tomarse a la ligera. Incluso algunos profesores piden a los alumnos que construyan instrumentos de medición de un día para otro, o lo que es casi lo mismo, de una semana a otra. Lo cual habla del poco o nulo conocimiento del proceso de elaboración de instrumentos de medición. Esta improvisación genera —casi siempre— instrumentos poco válidos o confiables y no debe existir en la investigación social (menos aún en ambientes académicos). Aun a los investigadores experimentados les toma tiempo desarrollar un instrumento de medición. Es por ello que los construyen con cuidado y frecuentemente están desarrollándolos, para que cuando los necesiten con premura se encuentren preparados para aplicarlos, pero no los improvisan. Además, para poder construir un instrumento de medición se requiere conocer muy bien a la variable que se pretende medir y la teoría que la sustenta. Por ejemplo, generar —o simplemente seleccionar— un instrumento que mida la inteligencia, la personalidad o los usos y gratificaciones de la televisión para el niño, requiere amplios conocimientos en la materia, estar actualizados al respecto y revisar cuidadosamente la literatura correspondiente.

El segundo factor es que a veces se utilizan instrumentos desarrollados en el extranjero que no han sido validados a nuestro contexto: cultura y tiempo. Traducir un instrumento —aun cuando adaptemos los términos a nuestro lenguaje y los contextualicemos— no es de ninguna manera (ni remotamente) validarlo. Es un primer y necesario paso, pero sólo es el principio. Por otra parte, hay instrumentos que fueron validados en nuestro contexto pero hace mucho tiempo. Hay instrumentos que hasta el lenguaje nos suena "arcaico". Las culturas, los grupos y las personas cambian; y esto debemos tomarlo en cuenta al elegir o desarrollar un instrumento de medición.

Un tercer factor es que en ocasiones el instrumento resulta inadecuado para las personas a las que se les aplica: no es empático. Utilizar un lenguaje muy elevado para el respondiente, no tomar en cuenta diferencias en cuanto a sexo, edad, conocimientos, capacidad de respuesta, memoria, nivel ocupacional y educativo, motivación para responder y otras diferencias en los respondientes; son errores que pueden afectar la validez y confiabilidad del instrumento de medición.

Un cuarto factor que puede influir está constituido por las condiciones en las que se aplica el instrumento de medición. Si hay ruido, hace mucho frío (por ejemplo

en una encuesta de casa en casa), el instrumento es demasiado largo o tedioso, son cuestiones que pueden afectar negativamente la validez y la confiabilidad. Normalmente en los experimentos se puede contar con instrumentos de medición más largos y complejos que en los diseños no experimentales. Por ejemplo, en una encuesta pública sería muy difícil poder aplicar una prueba larga o compleja.

Por otra parte, aspectos mecánicos tales como que si el instrumento es escrito, no se lean bien las instrucciones, falten páginas, no haya espacio adecuado para contestar, no se comprendan las instrucciones, también pueden influir de manera negativa.

9.4. ¿CÓMO SE SABE SI UN INSTRUMENTO DE MEDICIÓN ES CONFIABLE Y VÁLIDO?

En la práctica es casi imposible que una medición sea perfecta. Generalmente se tiene un grado de error. Desde luego, se trata de que este error sea el mínimo posible. Es por esto que la medición de cualquier fenómeno se conceptualiza con la siguiente fórmula básica:

$$X = t + e$$

Donde "X" representa los valores observados (resultados disponibles), "t" son los valores verdaderos y "e" es el grado de error en la medición. Si no hay error de medición ("e" es igual a cero), el valor observado y el verdadero son equivalentes. Esto puede verse claramente así:

$$X = t + 0$$

$$X = t$$

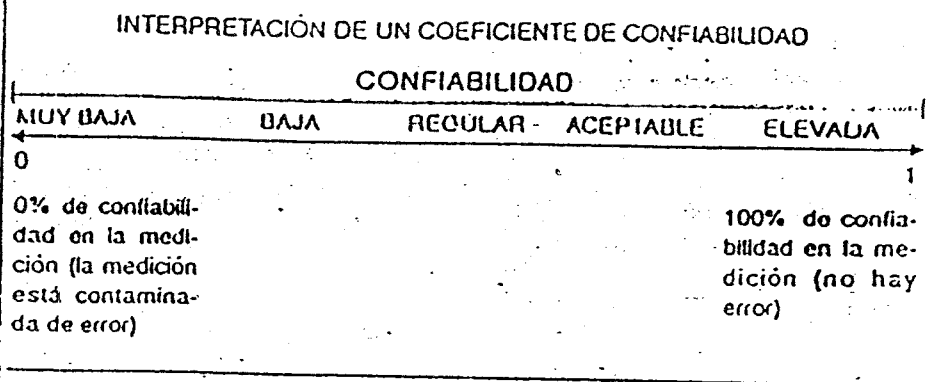
Esta situación representa el ideal de la medición. Entre mayor sea el error al medir, el valor que observamos (y que es en el que nos basamos) se aleja más del valor real o verdadero. Por ejemplo, si medimos la motivación de un individuo y esta medición está contaminada por un grado de error considerable, la motivación registrada por el instrumento será bastante diferente de la motivación real que tiene ese individuo. Por ello es importante que el error sea reducido lo más posible. Pero, ¿cómo sabemos el grado de error que tenemos en una medición? Calculando la confiabilidad y validez.

CÁLCULO DE LA CONFIABILIDAD

Existen diversos procedimientos para calcular la confiabilidad de un instrumento de medición. Todos utilizan fórmulas que producen coeficientes de confiabilidad. Estos coeficientes pueden oscilar entre 0 y 1. Donde un coeficiente de 0 significa nula confiabilidad y 1 representa un máximo de confiabilidad (confiabilidad total). Entre

mas se acerca el coeficiente a cero (0), hay mayor error en la medición. Esto se ilustra en la figura 9.4.

FIGURA 9.4



Los procedimientos más utilizados para determinar la confiabilidad mediante un coeficiente son:

1. *Medida de estabilidad (confiabilidad por test-retest).* En este procedimiento un mismo instrumento de medición (o ítems o indicadores)³⁰ es aplicado dos o más veces a un mismo grupo de personas, después de un período de tiempo. Si la correlación entre los resultados de las diferentes aplicaciones es altamente positiva, el instrumento se considera confiable. Se trata de una especie de diseño panel. Desde luego, el período de tiempo entre las mediciones es un factor a considerar. Si el período es largo y la variable susceptible de cambios, ello puede confundir la interpretación del coeficiente de confiabilidad obtenido por este procedimiento. Y si el período es corto las personas pueden recordar cómo contestaron en la primera aplicación del instrumento, para aparecer como más consistentes de lo que son en realidad (Bohmstedt, 1976).
2. *Método de formas alternativas o paralelas.* En este procedimiento no se administra el mismo instrumento de medición, sino dos o más versiones equivalentes de éste. Las versiones son similares en contenido, instrucciones, duración y otras características. Las versiones —generalmente dos— son administradas a un mismo grupo de personas dentro de un período de tiempo relativamente corto. El instrumento es confiable si la correlación entre los resultados de ambas administraciones es significativamente positiva. Los patrones de respuesta deben variar poco entre las aplicaciones.
3. *Método de mitales partidas (split-halves).* Los procedimientos anteriores (medida de estabilidad y método de formas alternas), requieren cuando menos dos administraciones de la medición en el mismo grupo de individuos. En cambio,

³⁰ Un ítem es la unidad mínima que compone a una medición; es un reactivo que estimula una respuesta en un sujeto (por ejemplo, una pregunta, una frase, una línea, fotografía, un objeto de descripción).

el método de mitades-partidas requiere sólo una aplicación de la medición. Específicamente, el conjunto total de ítems (o componentes) es dividido en dos mitades y las puntuaciones o resultados de ambas son comparados. Si el instrumento es confiable, las puntuaciones de ambas mitades deben estar fuertemente correlacionadas. Un individuo con baja puntuación en una mitad, tenderá a tener también una baja puntuación en la otra mitad. El procedimiento se diagrama en la figura 9.5.

FIGURA 9.5

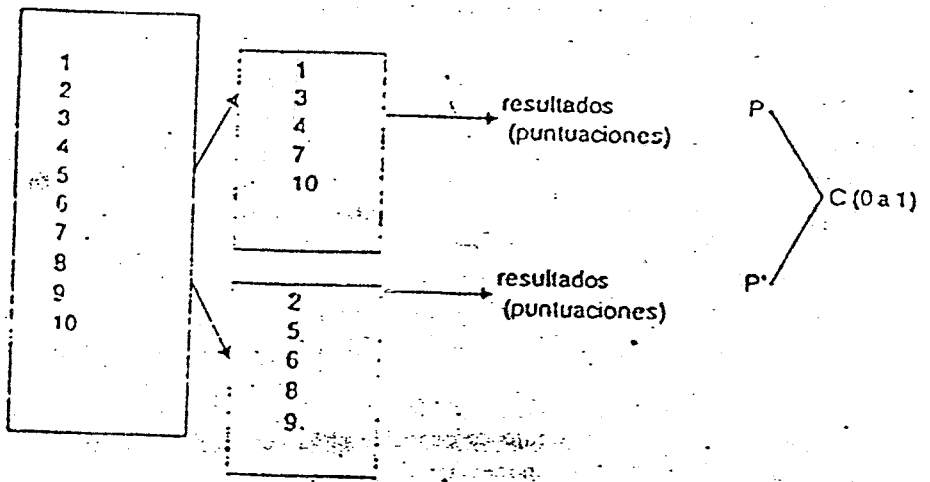
ESQUEMA DEL PROCEDIMIENTO DE MITADES-PARTIDAS

El instrumento de medición se aplica a un grupo

Los ítems se dividen en dos mitades (el instrumento se divide en dos)

Cada mitad se califica independientemente

Se correlacionan puntuaciones y se determina la confiabilidad



Al dividir los ítems, éstos se emparejan en contenido y dificultad.

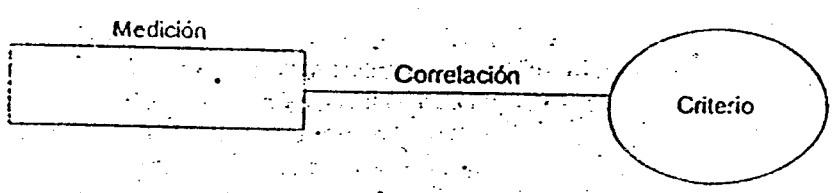
La confiabilidad varía de acuerdo al número de ítems que incluya el instrumento de medición. Cuantos más ítems la confiabilidad aumenta (desde luego, que se refieran a la misma variable). Esto resulta lógico, veámoslo con un ejemplo cotidiano: Si se desea probar qué tan confiable o consistente es la lealtad de un amigo hacia nuestra persona, cuantas más pruebas le pongamos, su confiabilidad será mayor. Claro está que demasiados ítems provocarán cansancio en el respondiente.

- 4. *Coefficiente alfa de Cronbach.* Este coeficiente desarrollado por J. L. Cronbach requiere una sola administración del instrumento de medición y produce valores que oscilan entre 0 y 1. Su ventaja reside en que no es necesario dividir en dos mitades a los ítems del instrumento de medición, simplemente se aplica la medición y se calcula el coeficiente. La manera de calcular este coeficiente se muestra en el siguiente capítulo.
- 5. *Coefficiente KR-20.* Kuder y Richardson (1937) desarrollaron un coeficiente para estimar la confiabilidad de una medición, su interpretación es la misma que la del coeficiente alfa.

CÁLCULO DE LA VALIDEZ

La validez de contenido es compleja de obtener. Primero, es necesario revisar cómo ha sido utilizada la variable por otros investigadores. Y en base a dicha revisión elaborar un universo de ítems posibles para medir la variable y sus dimensiones (el universo tiene que ser lo más exhaustivo que sea factible). Posteriormente, se consulta con investigadores familiarizados con la variable para ver si el universo es exhaustivo. Se seleccionan los ítems bajo una cuidadosa evaluación. Y si la variable tiene diversas dimensiones o facetas que la componen, se extrae una muestra probabilística de ítems (ya sea al azar o estratificada —cada dimensión constituiría un estrato—). Se administran los ítems, se correlacionan las puntuaciones de los ítems entre sí (debe haber correlaciones altas, especialmente entre ítems que miden una misma dimensión) (Bohmstedt, 1976), y se hacen estimaciones estadísticas para ver si la muestra es representativa. Para calcular la validez de contenido son necesarios varios coeficientes.

La validez de criterio es más sencilla de estimar, lo único que hace el investigador es correlacionar su medición con el criterio, y este coeficiente es el que se toma como coeficiente de validez (Bohmstedt, 1976). Esto podría representarse así: ³⁷



La validez de constructo se suele determinar mediante un procedimiento denominado "Análisis de Factores". Su aplicación requiere de sólidos conocimientos estadísticos y del uso de un programa estadístico apropiado en computadora. Para quien desee comprenderse con esta técnica recomendamos consultar a Harman (1967), Gorsuch (1974), Nie et al. (1975), On-Kim y Mueller (1978a y 1978b) y Hunter (1980). Asimismo, para aplicarlos se sugiere revisar a Nie et al. (1975), Cooper y Cenis (1976) y —en español— Padua (1979). Aunque es requisito conocer el programa estadístico para computadora. Esta técnica se describe en la página 420.

9.5. ¿QUÉ PROCEDIMIENTO SE SIGUE PARA CONSTRUIR UN INSTRUMENTO DE MEDICIÓN?

Existen diversos tipos de instrumentos de medición, cada uno con características diferentes. Sin embargo, el procedimiento general para construirlos es semejante. Antes de comentar este procedimiento, es necesario aclarar que en una investigación hay dos opciones respecto al instrumento de medición:

- 1) Elegir un instrumento ya desarrollado y disponible, el cual se adapta a los requerimientos del estudio en particular.
- 2) Construir un nuevo instrumento de medición de acuerdo con la técnica apropiada para ello.

En ambos casos es importante tener evidencia sobre la confiabilidad y validez del instrumento de medición.

El procedimiento que sugerimos para construir un instrumento de medición es el siguiente, especialmente para quien se inicia en esta materia.

PASOS

- a) LISTAR LAS VARIABLES que se pretende medir u observar.
- b) REVISAR SU DEFINICIÓN CONCEPTUAL Y COMPRENDER SU SIGNIFICADO. Por ejemplo, comprender bien qué es la motivación intrínseca y qué dimensiones la integran.
- c) REVISAR CÓMO HAN SIDO DEFINIDAS OPERACIONALMENTE LAS VARIABLES, esto es, cómo se ha medido cada variable. Ello implica comparar los distintos instrumentos o maneras utilizadas para medir las variables (comparar su confiabilidad, validez, sujetos a los cuales se les aplicó, facilidad de administración, veces que las mediciones han resultado exitosas y posibilidad de uso en el contexto de la investigación).
- d) ELEGIR EL INSTRUMENTO O LOS INSTRUMENTOS (YA DESARROLLADOS) QUE HAYAN SIDO FAVORECIDOS POR LA COMPARACIÓN Y ADAPTARLOS AL CONTEXTO DE LA INVESTIGACIÓN. En este caso sólo deben seleccionarse instrumentos cuya confiabilidad y validez se reporte. No se puede uno fiar de una manera de medir que carezca de evidencia clara y precisa de confiabilidad y validez. Cualquier investigación sería reportar la confiabilidad y validez de su instrumento de medición. Recuerdese que la primera vara de 0 a 1 y para la segunda se debe mencionar el método utilizado de validación y su interpretación. De no ser así no podemos asegurar que el instrumento sea el adecuado. Si se selecciona un instrumento desarrollado en otro país, deben hacerse pruebas piloto más extensas (véase el paso G). También, no debe olvidarse que traducir no es validar un instrumento, por muy buena que sea la traducción.

O en caso de que no se elija un instrumento ya desarrollado, sino que se prefiere construir o desarrollar uno propio, debe pensarse en cada variable y sus

dimensiones, y en indicadores precisos e ítems para cada dimensión. La figura 9.6 es un ejemplo de ello:

FIGURA 9.6

EJEMPLO DE DESARROLLO DE ÍTEMS

VARIABLE	DEFINICIÓN OPERACIONAL	INDICADORES	DIMENSIONES	ÍTEMS
Coordinación entre organizaciones compradoras y proveedoras, desde el punto de vista de las primeras.	Grado percibido mutuo de esfuerzo invertido para no provocar problemas a la otra parte al interferir en sus deberes y responsabilidades.	Grado percibido mutuo de interés y buena voluntad de ambas partes.	Coordinación de conflictos.	<p>¿Qué tanto se esfuerza su empresa por no provocar problemas con sus proveedores?</p> <p>1. Se esfuerza al mínimo posible.</p> <p>2. Se esfuerza poco.</p> <p>3. Se esfuerza medianamente.</p> <p>4. Se esfuerza mucho.</p> <p>5. Se esfuerza al máximo posible.</p>
			Coordinación de no interferencia.	<p>¿Qué tanto se esfuerzan sus proveedores por no provocar problemas con su empresa?</p> <p>1. Se esfuerzan al mínimo posible.</p> <p>2. Se esfuerzan poco.</p> <p>3. Se esfuerzan medianamente.</p> <p>4. Se esfuerzan mucho.</p> <p>5. Se esfuerzan al máximo posible.</p>
				<p>¿Cuánto se esfuerza su empresa por no interferir en los deberes y responsabilidades de sus proveedores?</p> <p>5. Se esfuerza al máximo posible.</p> <p>4. Se esfuerza mucho.</p>

VARIABLE	DEFINICIÓN OPERACIONAL	INDICADORES	DIMENSIONES	ÍTEMS
			Coordinación de objetivos.	<p>3. Se esfuerzan medianamente.</p> <p>2. Se esfuerzan poco.</p> <p>1. Se esfuerza al mínimo posible.</p> <p>¿Cuánto se esfuerzan sus proveedores por no interferir con los deberes y responsabilidades de su empresa?</p> <p>5. Se esfuerzan al máximo posible.</p> <p>4. Se esfuerzan mucho.</p> <p>3. Se esfuerzan medianamente.</p> <p>2. Se esfuerzan poco.</p> <p>1. Se esfuerzan al mínimo posible.</p> <p>¿Cuánto se esfuerza la empresa por trabajar junto con sus proveedores —de manera constante— para alcanzar objetivos comunes?</p> <p>5. Se esfuerza al máximo posible.</p> <p>4. Se esfuerza mucho.</p> <p>3. Se esfuerza medianamente.</p> <p>2. Se esfuerza poco.</p> <p>1. Se esfuerza al mínimo posible.</p> <p>¿Cuánto se esfuerzan los proveedores por trabajar junto con su empresa —de manera constante— para alcanzar objetivos comunes?</p>

VARIABLE	DEFINICIÓN OPERACIONAL	INDICADORES	DIMENSIONES	ÍTEMES
			Coordinación de objetivos.	5. Se esfuerzan al máximo posible. 4. Se esfuerzan mucho. 3. Se esfuerzan medianamente. 2. Se esfuerzan poco. 1. Se esfuerzan al mínimo posible.
			Coordinación de rutinas.	En general, ¿qué tan bien establecidas están las rutinas para el trato de la empresa con sus proveedores? 5. Muy bien establecidas. 4. Bien establecidas. 3. Medianamente establecidas. 2. Mal establecidas. 1. Muy mal establecidas.
Frecuencia de la interacción entre organizaciones.	Lapsos de interacciones entre organizaciones.	Lapso máximo entre interacciones de comunicación.	Visitas de representantes.	Estableciendo un promedio aproximado ¿qué tan seguido recibe su empresa la visita de los representantes de sus proveedores verdaderamente importantes? 13. Varias veces al día. 12. Una vez al día. 11. Tres veces por semana. 10. Dos veces por semana. 9. Una vez a la semana. 8. Tres veces al mes. 7. Dos veces al mes. 6. Una vez al mes.

VARIABLE	DEFINICIÓN OPERACIONAL	INDICADORES	DIMENSIONES	ÍTEMS
				5. Una vez cada dos meses. 4. Una vez cada cuatro meses. 3. Una vez cada seis meses. 2. Una vez al año. 1. Otra (especifique). Estableciendo un promedio aproximado ¿qué tan seguido recibe su empresa la visita de los representantes de sus proveedores poco importantes? 13. Varias veces al día. 12. Una vez al día. 11. Tres veces por semana. 10. Dos veces por semana. 9. Una vez a la semana. 8. Tres veces al mes. 7. Dos veces al mes. 6. Una vez al mes. 5. Una vez cada dos meses. 4. Una vez cada cuatro meses. 3. Una vez cada seis meses. 2. Una vez al año. 1. Otra (especifique).
			Llamadas telefónicas.	Estableciendo un promedio aproximado ¿qué tan seguido le llaman por teléfono a su empresa los representantes de sus proveedores muy importantes?

VARIABLE	DEFINICIÓN OPERACIONAL	INDICADORES	DIMENSIONES	ÍTEMS
				13. Varias veces al día.
				12. Una vez al día.
				11. Tres veces por semana.
				10. Dos veces por semana.
				9. Una vez a la semana.
				8. Tres veces al mes.
				7. Dos veces al mes.
				6. Una vez al mes.
				5. Una vez cada dos meses.
				4. Una vez cada cuatro meses.
				3. Una vez cada seis meses.
				2. Una vez al año.
				1. Otra (especifique).
				Etcétera.

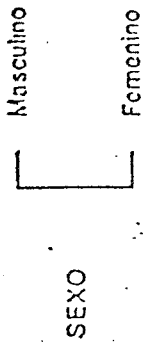
En este segundo caso, debemos asegurarnos de tener un suficiente número de ítems para medir todas las variables en todas sus dimensiones. Ya sea que se seleccione un instrumento previamente desarrollado y se adapte o bien, se construya uno, éste constituye la versión preliminar de nuestra medición. Versión que debe pulirse y ajustarse, como se verá más adelante.

c) **INDICAR EL NIVEL DE MEDICIÓN DE CADA ÍTEM Y, por ende, EL DE LAS VARIABLES.**

Existen CUATRO NIVELES DE MEDICIÓN ampliamente conocidos:

1. *Nivel de medición nominal.* En este nivel se tienen dos o más categorías del ítem o variable. Las categorías no tienen orden o jerarquía. Lo que se mide es colocado en una u otra categoría, lo que indica solamente diferencias respecto a una o más características.

Por ejemplo, la variable sexo de la persona tiene sólo dos categorías: masculino y femenino (si la variable fuera "práctica sexual" podría haber tal vez más, pero sexo sólo tiene dos categorías). Ninguna de las categorías tiene mayor jerarquía que la otra, las categorías únicamente reflejan diferencias en la variable. No hay orden de mayor a menor.

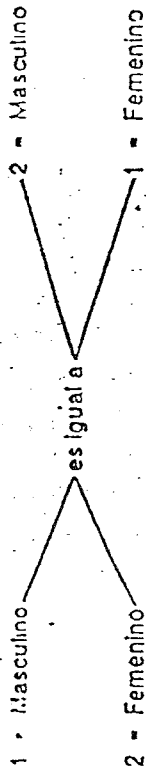


Si les asignamos una etiqueta o símbolo a cada categoría, éste exclusivamente define a la categoría. Por ejemplo:

• = Masculino

z = Femenino

Si usamos numerales es lo mismo:



Los números utilizados en este nivel de medición tienen una función puramente clasificatoria y no se pueden manipular aritméticamente. Por ejemplo, la afiliación religiosa es una variable nominal, si pretendiéramos operarla aritméticamente tendríamos situaciones tan ridículas como ésta:

- 1 = Católico
- 2 = Judío
- 3 = Protestante
- 4 = Musulmán
- 5 = Otros

1 + 2 = 3

Un católico + un judío = protestante?

No tiene sentido.

Las variables nominales pueden incluir dos categorías (se las llama dicotómicas) o más categorías (se las llama categóricas). Ejemplos de variables nominales serían el sexo y el tipo de escuela a la que se asiste (privada-pública); y de variables categóricas tendríamos a la afiliación política (Partido A, Partido B, ...), la raza elegida, la raza, el departamento o provincia o estado de nacimiento y el canal de televisión preferido.

Nivel de medición ordinal. En este nivel se tienen varias categorías, pero además éstas mantienen un orden de mayor a menor. Las etiquetas o símbolos de las categorías si indican jerarquía. Por ejemplo, el prestigio ocupacional en los Estados Unidos ha sido medido por diversas escalas que ordenan a las profesiones de acuerdo con su prestigio, por ejemplo:

90	Ingeniero químico.
80	Científico de ciencias naturales (excluyendo la Química).
60	Actor
50	Operador de estaciones eléctricas de potencia.
02	Manufactureros de tabaco.

90 es más que 80, 80 más que 60, 60 más que 50 y así sucesivamente (los números —símbolos de categorías— definen posiciones). Sin embargo, las categorías no están ubicadas a intervalos iguales (no hay un intervalo común). No podríamos decir con exactitud que entre un actor (60) y un operador de estaciones de poder (50) existe la misma distancia —en prestigio— que entre un científico de las ciencias naturales (80) y un ingeniero químico (90). Aparentemente en ambos casos la distancia es 10, pero no es una distancia real. Otra escala "clasificó el prestigio de dichas profesiones de la siguiente manera:

Valor en la escala	Profesión
98	Ingeniero químico.
95	Científico de ciencias naturales (excluyendo la Química).
84	Actor.
78	Operador de estaciones eléctricas de potencia.
13	Manufactureros de tabaco.

Aquí la distancia entre un actor (84) y un operador de estaciones (78) es de 6, y la distancia entre un ingeniero químico (98) y un científico de ciencias naturales (95) es de 3.

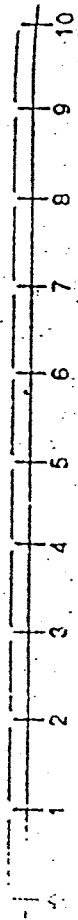
Otro ejemplo sería la posición jerárquica en la empresa:

Presidente	10
Vicepresidente	9
Director General	8
Gerente de Área	7
Subgerente o Superintendente	6
Jefe	5
Empleado A	4
Empleado B	3
Empleado C	2
Indefinición	1

Sabemos que el Presidente (10) es más que el Vicepresidente (9), éste más que el Director General (8), a su vez este último más que el Gerente (7) y así sucesivamente; pero no puede precisarse en cada caso cuánto más. Tampoco podemos utilizar las operaciones aritméticas básicas; No podríamos decir que 4 (empleando A) + 5 (jefe) = 9 (Vicepresidente), ni que 10 (Presidente) + 5 (jefe) = 2 (empleando C). Sería absurdo, no tiene sentido.

3. *Nivel de medición por intervalos.* Además de haber orden o jerarquía entre categorías, se establecen intervalos iguales en la medición. Las distancias entre categorías son las mismas a lo largo de toda la escala. Hay intervalo constante, una unidad de medida.

Intervalo constante



Por ejemplo: Una prueba de resolución de problemas matemáticos (30 problemas de igual dificultad). Si Ana Cecilia resolvió 10, Laura resolvió 20 y Brenda 30. La distancia entre Ana Cecilia y Laura es igual a la distancia entre Laura y Brenda.

Sin embargo, el cero (0) en la medición, es un cero arbitrario, no es real (se asigna arbitrariamente a una categoría, el valor de cero y a partir de ésta se construye la escala). Un ejemplo clásico en ciencias naturales es la temperatura (en grados centígrados y Fahrenheit), el cero es arbitrario, no implica que realmente haya cero temperatura (incluso en ambas escalas el cero es diferente).

Cabe agregar que diversas mediciones en el estudio del comportamiento humano no son verdaderamente de intervalo (v.g. escalas de actitudes, pruebas de inteligencia y de otros tipos), pero se acercan a este nivel y se suele tratarlas como si fueran mediciones de intervalo. Esto se hace porque este nivel de medición permite utilizar las operaciones aritméticas básicas (suma, resta, multiplicación y división) y algunas estadísticas medidas, que de otro modo no se usaban. Aunque algunos investigadores no están de acuerdo en suponer tales mediciones como si fueran de intervalo (pero estos investigadores son minoría).

3. *Nivel de medición por razones.* En este nivel, además de tenerse todas las características del nivel de intervalos (intervalos iguales entre las categorías y aplicación de operaciones aritméticas básicas y sus derivaciones), el cero es real, es absoluto (no es arbitrario). Cero absoluto implica que hay un punto en la escala donde no existe la propiedad.



Ejemplos de estas mediciones serían la exposición a la televisión, el número de hijos, la productividad, las ventas de un producto y el ingreso.

Desde luego, hay variables que pueden medirse en más de un nivel, según el propósito de medición. Por ejemplo, la variable "antigüedad en la empresa":

Nivel de medición	Categorías
— De razón	En días (0 a X días)
— Ordinal	Bastante antigüedad Regular antigüedad Poca antigüedad

Es muy importante indicar el nivel de medición de todas las variables e ítems de la investigación, porque dependiendo de dicho nivel se selecciona uno u otro tipo de análisis estadístico (por ejemplo, la prueba estadística para correlacionar dos variables de intervalo es muy distinta a la prueba para correlacionar dos variables ordinales). Así, es necesario hacer una relación de variables, ítems y niveles de medición.

INDICAR LA MANERA COMO SE HABRÁN DE CODIFICAR LOS DATOS en cada ítem y variable. CODIFICAR los datos significa asignarles un valor numérico que los represente. Es decir, a las categorías de cada ítem y variable se les asignan valores numéricos que tienen un significado. Por ejemplo, si tuviéramos la variable "sexo" con sus respectivas categorías, "masculino" y "femenino", a cada categoría le asignaríamos un valor. Este podría ser:

Categoría	Codificación (valor asignado)
— Masculino	1
— Femenino	2

Así, Carla Magaña en la variable sexo sería un "2". Luis Gerardo Vera y Rubén Reyes serían un "1". Verónica Laríos un "2" y así sucesivamente.

Otro ejemplo sería la variable "horas de exposición diaria a la televisión", que podría codificarse de la siguiente manera:

Categoría	Codificación (valor asignado)
— No ve televisión	0
— Menos de una hora	1
— Una hora	2
— Más de una hora, pero menos de dos	3
— Dos horas	4
— Más de dos horas, pero menos de tres	5

— Tres horas	6
— Más de tres horas, pero menos de cuatro	7
— Cuatro horas	8
— Más de cuatro horas	9

Es necesario insistir que cada ítem y variable deberán tener una *codificación* (códigos numéricos) para sus categorías. Desde luego, hay veces que un ítem no puede ser codificado a priori (precodificado) porque es sumamente difícil conocer cuáles serán sus categorías. Por ejemplo, si en una investigación fuéramos a preguntar: ¿Qué opina del programa económico que recientemente aplicó el Gobierno? Es posible que las categorías encontradas podrían ser muchas más de las que nos imaginemos y resultaría difícil predecir con precisión cuántas y cuáles serán. En estos casos la codificación se lleva a cabo una vez que se aplica el ítem (*a posteriori*). A lo largo de este capítulo se profundizará en la forma de codificar y sus implicaciones. Por el momento, lo importante es que se comprenda el significado de codificar y que el instrumento de medición, antes de aplicarse, debe ir precodificado hasta donde sea posible (codificar los ítems cuyas categorías sean conocidas de antemano).

La codificación es necesaria para poder cuantitativamente analizar los datos (aplicar análisis estadístico). A veces se utilizan letras o símbolos en lugar de números (*, A, Z).

g) UNA VEZ QUE SE INDICA EL NIVEL DE MEDICIÓN DE CADA VARIABLE E ÍTEM Y QUE SE DETERMINA SU CODIFICACIÓN, SE PROCEDE A APLICAR UNA "PRUEBA PILOTO" DEL INSTRUMENTO DE MEDICIÓN. Es decir, se aplica a personas con características semejantes a las de la muestra o población objetivo de la investigación.

En esta prueba se analiza si las instrucciones se comprenden y si los ítems funcionan adecuadamente. Los resultados se usan para calcular la confiabilidad —y de ser posible la validez— del instrumento de medición.

La *prueba piloto* se realiza con una pequeña muestra (inferior a la muestra definitiva). Los autores aconsejamos que cuando la muestra sea de 200 o más, se lleve a cabo la prueba piloto con entre 25 y 60 personas. Salvo que la investigación exija un número mayor.

h) SOBRE LA BASE DE LA PRUEBA PILOTO, EL INSTRUMENTO DE MEDICIÓN PRELIMINAR SE MODIFICA, AJUSTA Y SE MEJORA. LOS INDICADORES DE CONFIABILIDAD Y VALIDEZ SON UNA BUENA AYUDA. Y ESTAREMOS EN CONDICIONES DE APLICARLO.

Este procedimiento general para desarrollar una medición debe —desde luego— adaptarse a las características de los diferentes tipos de instrumentos de que disponemos en el estudio del comportamiento, los cuales veremos a continuación.

9.6. ¿DE QUÉ TIPOS DE INSTRUMENTOS DE MEDICIÓN O RECOLECCIÓN DE LOS DATOS DISPONEMOS EN LA INVESTIGACIÓN SOCIAL?

En la investigación del comportamiento disponemos de diversos tipos de instrumentos para medir las variables de interés y en algunos casos se pueden combinar dos o más métodos de recolección de los datos. A continuación describimos —brevemente— estos métodos o tipos de instrumentos de medición.

9.6.1. Escalas para medir las actitudes

Una *actitud* es una predisposición aprendida para responder consistentemente de una manera favorable o desfavorable respecto a un objeto o sus símbolos (Fishbein y Ajzen, 1975; Oskamp, 1977). Así, los seres humanos tenemos actitudes hacia muy diversos objetos o símbolos, por ejemplo: actitudes hacia el aborto, la política económica, la familia, un profesor, diferentes grupos étnicos, la Ley, nuestro trabajo, el nacionalismo, hacia nosotros mismos, etcétera.

Las actitudes están relacionadas con el comportamiento que mantenemos en torno a los objetos a que hacen referencia. Si mi actitud hacia el aborto es desfavorable, probablemente no abortaría o no participaría en un aborto. Si mi actitud es favorable a un partido político, lo más probable es que vote por él en las próximas elecciones. Desde luego, las actitudes sólo son un indicador de la conducta, pero no la conducta en sí. Es por ello que las mediciones de actitudes deben interpretarse como "síntomas" y no como "hechos" (Padua, 1979). Por ejemplo, si detecto que la actitud de un grupo hacia la contaminación es desfavorable, esto no significa que las personas están adoptando acciones para evitar contaminar el ambiente, pero sí es un indicador de que pueden ir las adoptando paulatinamente. La actitud es como una "semilla", que bajo ciertas condiciones puede "germinar en comportamiento".

Las actitudes tienen diversas propiedades, entre las que destacan: dirección (positiva o negativa) e intensidad (alta o baja), estas propiedades forman parte de la medición.

Los métodos más conocidos para medir por escalas las variables que constituyen actitudes son: el método de escalamiento Likert, el diferencial semántico y la escala de Guttman. Hablaremos de cada método.

Escalamiento tipo Likert⁴⁰

Este método fue desarrollado por Rensis Likert a principios de los treinta; sin embargo, se trata de un enfoque vigente y bastante popularizado. Consiste en un conjunto de ítems presentados en forma de afirmaciones o juicios ante los cuales se pide la reacción de los sujetos a los que se les administra. Es decir, se presenta cada afirmación y se

⁴⁰ Para profundizar en esta técnica se recomienda consultar a Likert (1976a u 1976b), Seiler y Hough (1976) y Padua (1979).

pide al sujeto que exprese su reacción eligiendo uno de los cinco puntos de la escala. A cada punto se le asigna un valor numérico. Así, el sujeto obtiene una puntuación respecto a la afirmación y al final se obtiene su puntuación total sumando las puntuaciones obtenidas en relación a todas las afirmaciones.

Las afirmaciones califican al objeto de actitud que se está midiendo y deben expresar sólo una relación lógica, además es muy recomendable que no excedan de —aproximadamente— 20 palabras.

EJEMPLO

Objeto de actitud medido	Afirmación
El voto	"Votar es una obligación de todo ciudadano responsable"

En este caso la afirmación incluye 8 palabras y expresa una sola relación lógica (X—Y). Las alternativas de respuesta o puntos de la escala son cinco e indican cuánto se está de acuerdo con la afirmación correspondiente. Las alternativas más comunes se presentan en la figura 9.7. Debe recordarse que a cada una de ellas se le asigna un valor numérico y sólo puede marcarse una opción. Se considera un dato inválido a quien marque dos o más opciones.

FIGURA 9.7

ALTERNATIVAS O PUNTOS EN LAS ESCALAS LIKERT

Alternativa 1:

		"Afirmación"		
Muy de acuerdo	De acuerdo	Ni de acuerdo, ni en desacuerdo	En desacuerdo	Muy en desacuerdo

Alternativa 2:

		"Afirmación"		
Totalmente de acuerdo	De acuerdo	Neutral	En desacuerdo	Totalmente en desacuerdo

Alternativa 3:

		"Afirmación"		
Definitivamente sí	Probablemente sí	Indeciso	Probablemente no	Definitivamente no

Alternativa 4:

		"Afirmación"		
Completamente verdadero	Verdadero	Ni falso, ni verdadero	Falso	Completamente falso

Asimismo, pueden hacerse distintas combinaciones como "totalmente verdadero" o "completamente no". Y las alternativas de respuesta pueden colocarse horizontalmente —como en la figura 9.7— o verticalmente.

EJEMPLO

- () Muy de acuerdo
- () De acuerdo
- () Ni de acuerdo, ni en desacuerdo
- () En desacuerdo
- () Muy en desacuerdo

O bien utilizando recuadros en lugar de parentesis:

- Definitivamente si
- Probablemente si
- Indeciso
- Probablemente no
- Definitivamente no

Es indispensable comentar que el número de categorías de respuesta debe ser el mismo para todas las afirmaciones.

DIRECCIÓN DE LAS AFIRMACIONES

Las afirmaciones pueden tener dirección: *favorable o positiva y desfavorable o negativa*. Y esta dirección es muy importante para saber cómo se codifican las alternativas de respuesta.

Si la afirmación es positiva significa que califica favorablemente al objeto de actitud, y entre los sujetos estén más de acuerdo con la afirmación, su actitud es más favorable.

EJEMPLO

"El Ministerio de Hacienda ayuda al contribuyente a resolver sus problemas en el pago de impuestos".

Si estamos "muy de acuerdo" implica una actitud más favorable hacia el Ministerio de Hacienda que si estamos "de acuerdo". En cambio, si estamos "muy en desacuerdo" implica una actitud muy desfavorable. Por lo tanto, cuando las afirmaciones son positivas se califican comúnmente de la siguiente manera:

- (5) Muy de acuerdo
- (4) De acuerdo
- (3) Ni de acuerdo, ni en desacuerdo
- (2) En desacuerdo
- (1) Muy en desacuerdo

Es decir, estar más de acuerdo implica una puntuación mayor.

Si la afirmación es negativa significa que califica desfavorablemente al objeto de actitud, y entre los sujetos estén más de acuerdo con la afirmación, su actitud es menos favorable, esto es, más desfavorable.

EJEMPLO

"El Ministerio de Hacienda se caracteriza por obstaculizar al contribuyente en el pago de impuestos".

Si estamos "muy de acuerdo" implica una actitud más desfavorable que si estamos de "acuerdo" y así sucesivamente. En contraste, si estamos "muy en desacuerdo" implica una actitud favorable hacia el Ministerio de Hacienda. Rechazamos la frase porque califica negativamente al objeto de actitud. Un ejemplo cotidiano de afirmación negativa sería: "Luis es un mal amigo", entre más de acuerdo estemos con la afirmación, nuestra actitud hacia Luis es menos favorable. Es decir, estar más de acuerdo implica una puntuación menor. Cuando las afirmaciones son negativas se califican al contrario de las positivas.

EJEMPLO

- (1) Totalmente de acuerdo
- (2) De acuerdo
- (3) Ni de acuerdo, ni en desacuerdo
- (4) En desacuerdo
- (5) Totalmente en desacuerdo

En la figura 9.8. se presenta un ejemplo de una escala Likert para medir la actitud hacia un organismo tributario.⁴¹

FIGURA 9.8

EJEMPLO DE UNA ESCALA LIKERT

LAS AFIRMACIONES QUE VOY A LEERLE SON OPINIONES CON LAS QUE ALGUNAS PERSONAS ESTÁN DE ACUERDO Y OTRAS EN DESACUERDO. VOY A PEDIRLE QUE ME DIGA POR FAVOR QUE TAN DE ACUERDO ESTÁ USTED CON CADA UNA DE ESTAS OPINIONES.

⁴¹ El ejemplo fue utilizado en un país latinoamericano y su confiabilidad total fue de 0.89, aquí se presenta una versión reducida de la escala original. El nombre del organismo tributario que aquí se utiliza es ficticio.

1. "EL PERSONAL DE LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES ES GROSERO AL ATENDER AL PÚBLICO".
1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo

2. "LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES SE CARACTERIZA POR LA DESHONESTIDAD DE SUS FUNCIONARIOS".
1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo

3. "LOS SERVICIOS QUE PRESTA LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES SON EN GENERAL MUY BUENOS".
5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo

4. "LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES INFORMA CLARAMENTE SOBRE CÓMO, DÓNDE Y CUÁNDO PAGAR LOS IMPUESTOS".
5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo

5. "LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES ES MUY LENTA EN LA DEVOLUCIÓN DE IMPUESTOS PAGADOS EN EXCESO".
1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo

6. "LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES INFORMA OPORTUNAMENTE SOBRE CÓMO, DÓNDE Y CUÁNDO PAGAR LOS IMPUESTOS".
5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo

7. "LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES TIENE NORMAS Y PROCEDIMIENTOS BIEN DEFINIDOS PARA EL PAGO DE IMPUESTOS".
5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo

8. "LA DIRECCIÓN GENERAL DE IMPUESTOS NACIONALES TIENE MALAS RELACIONES CON LA GENTE PORQUE COBRA IMPUESTOS MUY ALTOS".
1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo

Como puede observarse en la figura 9.8, las afirmaciones 1, 2, 5 y 8 son negativas (desfavorables) y las afirmaciones 3, 4, 6 y 7 son positivas (favorables).

FORMA DE OBTENER LAS PUNTUACIONES

Las puntuaciones de las escalas Likert se obtienen sumando los valores obtenidos respecto a cada frase. Por ello se le denomina escala aditiva. La figura 9.9 constituiria un ejemplo de cómo calificar una escala de Likert:

FIGURA 9.9

EJEMPLO DE CÓMO CALIFICAR UNA ESCALA LIKERT.

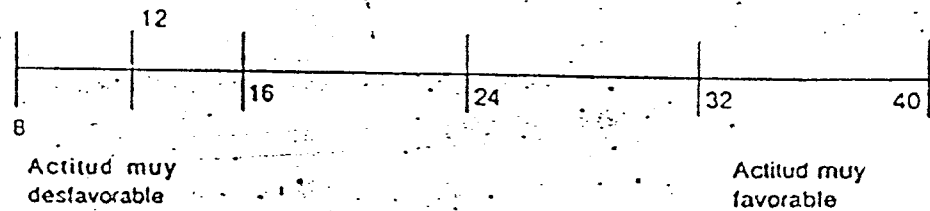
- 1. El personal de la dirección general de impuestos nacionales es grosero al atender al público.
 1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
 2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo
- 2. La dirección general de impuestos nacionales se caracteriza por la deshonestidad de sus funcionarios.
 1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
 2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo
- 3. Los servicios que presta la dirección general de impuestos nacionales son en general muy buenos.
 5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
 4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo
- 4. La dirección general de impuestos nacionales informa claramente sobre cómo, dónde y cuándo pagar los impuestos.
 5) Muy de acuerdo Ni de acuerdo, ni en desacuerdo
 4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo
- 5. La dirección general de impuestos nacionales es muy lenta en la devolución de impuestos pagados en exceso.
 1) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
 2) De acuerdo 4) En desacuerdo 5) Muy en desacuerdo
- 6. La dirección general de impuestos nacionales informa oportunamente sobre cómo, dónde y cuándo pagar los impuestos.
 5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
 4) De acuerdo 2) En desacuerdo 1) Muy en desacuerdo
- 7. La dirección general de impuestos nacionales tiene normas y procedimientos bien definidos para el pago de impuestos.
 5) Muy de acuerdo 3) Ni de acuerdo, ni en desacuerdo
 4) De acuerdo En desacuerdo 1) Muy en desacuerdo

8. La dirección general de impuestos nacionales tiene malas relaciones con la gente porque cobra impuestos muy altos.

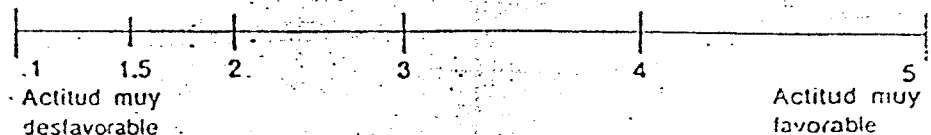
- 1) Muy de acuerdo
- 2) De acuerdo
- 3) Ni de acuerdo ni en desacuerdo
- 4) En desacuerdo
- 5) Muy en desacuerdo

VALOR = 1 + 2 + 1 + 3 + 1 + 1 + 2 + 1 = 12

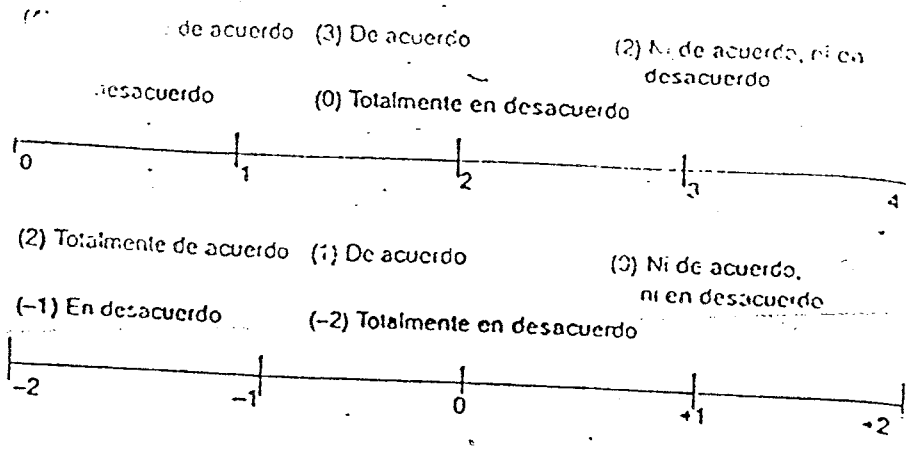
Una puntuación se considera alta o baja según el número de ítems o afirmaciones. Por ejemplo, en la escala para evaluar la actitud hacia el organismo tributario la puntuación mínima posible es de 8 (1+1+1+1+1+1+1) y la máxima es de 40 (5+5+5+5+5+5+5), porque hay ocho afirmaciones. La persona del ejemplo obtuvo "12", su actitud hacia el organismo tributario es más bien sumamente desfavorable, veámoslo gráficamente:



Si alguien hubiera tenido una puntuación de 37 (5+5+4+5+5+4+4+5) su actitud puede calificarse como sumamente favorable. En las escalas Likert a veces se califica el promedio obtenido en la escala mediante la sencilla fórmula $\frac{PT}{NT}$ (donde PT es la puntuación total en la escala y NT es el número de afirmaciones), y entonces una puntuación se analiza en el continuo 1:5 de la siguiente manera, con el ejemplo de quien obtuvo 12 en la escala ($\frac{12}{8} = 1.5$):



La escala Likert es, en estricto sentido, una medición ordinal, sin embargo, es común que se le trabaje como si fuera de intervalo. Asimismo, a veces se utiliza un rango de 0 a 4 o de -2 a +2 en lugar de 1 a 5. Pero esto no importa porque se cambia el marco de referencia de la interpretación. Veámoslo gráficamente



00134

Simply se ajusta el marco de referencia, pero el rango se mantiene y las categorías continúan siendo cinco.

OTRAS CONSIDERACIONES SOBRE LA ESCALA LIKERT

A veces se acorta o incrementa el número de categorías, sobre todo cuando los respondientes potenciales pueden tener una capacidad muy limitada de discriminación o por el contrario muy amplia.

EJEMPLOS (CON AFIRMACIONES POSITIVAS)

(1) De acuerdo	(0) En desacuerdo	
(3) De acuerdo	(2) Ni de acuerdo, ni en desacuerdo	(1) En desacuerdo
(7) Totalmente de acuerdo	(6) De acuerdo	(5) Indeciso, pero más bien de acuerdo
(4) Indeciso, ni de acuerdo ni en desacuerdo	(3) Indeciso, pero más bien en desacuerdo	
(2) En desacuerdo	(1) Totalmente en desacuerdo	

Si los respondientes tienen poca capacidad de discriminar pueden incluirse dos o tres categorías. Por el contrario, si son personas con un nivel educativo elevado y capacidad de discriminación, pueden incluirse siete categorías. Pero debe recalarse que el número de categorías de respuesta debe ser el mismo para todos los ítems, si son tres, son tres categorías para todos los ítems o afirmaciones. Si son cinco, son cinco categorías para todos los ítems.

Un aspecto muy importante de la escala Likert es que asume que los ítems o afirmaciones miden la actitud hacia un único concepto subyacente, si se van a medir actitudes hacia varios objetos, deberá incluirse una escala por objeto aunque se presenten conjuntamente, pero se califican por separado. En cada escala se considera que todos los ítems tienen igual peso.

CÓMO SE CONSTRUYE UNA ESCALA LIKERT

En términos generales, una escala Likert se construye generando un elevado número de afirmaciones que califiquen el objeto de actitud y se administran a un grupo piloto para obtener las puntuaciones del grupo en cada afirmación. Estas puntuaciones se correlacionan con las puntuaciones del grupo a toda la escala (la suma de las puntuaciones de todas las afirmaciones), y las afirmaciones cuyas puntuaciones se correlacionen significativamente con las puntuaciones de toda la escala, se seleccionan para integrar el instrumento de medición. Asimismo, debe calcularse la confiabilidad y validez de la escala.

PREGUNTAS EN LUGAR DE AFIRMACIONES

En la actualidad, la escala original se ha extendido a preguntas y observaciones. Por ejemplo:

¿CÓMO CONSIDERA USTED AL CONDUCTOR QUE APARECE EN LOS PROGRAMAS?

5 Muy buen conductor 4 Buen conductor 3 Regular
2 Mal conductor 1 Muy mal conductor

Esta pregunta se hizo como parte de la evaluación de un video empresarial. Otro ejemplo sería una pregunta que se hizo en una investigación para analizar la relación de compra-venta en empresas de la Ciudad de México (Paniagua, 1986). De ella se presenta un fragmento en la figura 9.10.

FIGURA 9.10.

EJEMPLO DE LA ESCALA LIKERT APLICADA A UNA PREGUNTA PARA ELEGIR SUS PROVEEDORES. ¿QUÉ TAN IMPORTANTE ES...?

	Indispensable (5)	Sumamente importante (4)	Medianamente importante (3)	Poco importante (2)	No se toma en cuenta (1)
- Precio					
- Forma de pago (contado-crédito)					
- Tiempo de entrega					
- Costo de entrega					

	Indispensable (5)	Sumamente importante (4)	Medianamente importante (3)	Poco importante (2)	No se toma en cuenta (1)
- Garantía del producto					
- Servicio de reparación					
- Prestigio del producto (marca)					
- Prestigio de la empresa proveedora					
- Comunicación que se tiene con la(s) persona(s) que representan al proveedor					
- Apego del proveedor a los requerimientos legales del producto					
- Cumplimiento del proveedor con las especificaciones					
- Información que sobre el producto proporcione el proveedor					
- Tiempo de trabajar con el proveedor					
- Entrega del producto en las condiciones acordadas					
- Calidad del producto					
- Personalidad de los vendedores					

Las respuestas se califican del mismo modo que ya hemos comentado.

Existen dos formas básicas de aplicar una escala Likert. La primera es de manera *autoadministrada*; se le entrega la escala al respondiente y éste marca respecto a cada afirmación, la categoría que mejor describe su reacción o respuesta. Es decir, marcan su respuesta. La segunda forma es *la entrevista*; un entrevistador lee las afirmaciones y alternativas de respuesta al sujeto y anota lo que éste conteste. Cuando se aplica vía entrevista, es muy necesario que se le entregue al respondiente una tarjeta donde se muestran las alternativas de respuesta o categorías. El siguiente es un ejemplo que se aplica a la pregunta de la figura 9.10:

Indispensable	Sumamente importante	Medianamente importante	Poco importante	No es tema en cuenta
---------------	----------------------	-------------------------	-----------------	----------------------

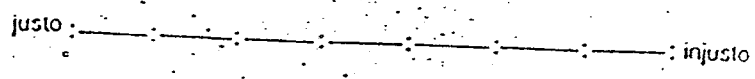
Al construir una escala Likert debemos asegurar que las afirmaciones y alternativas de respuesta serán comprendidas por los sujetos a los que se les aplicará y que éstos tendrán la capacidad de discriminación requerida. Ello se evalúa cuidadosamente en la prueba piloto.

*Diferencial semántico*⁴²

El diferencial semántico fue desarrollado originalmente por *Osgood, Suci y Tannenbaum* (1957) para explorar las dimensiones del significado. Pero hoy en día consiste en una serie de adjetivos extremos que califican al objeto de actitud, ante los cuales se solicita la reacción del sujeto. Es decir, éste tiene que calificar al objeto de actitud en un conjunto de adjetivos bipolares, entre cada par de adjetivos se presentan varias opciones y el sujeto selecciona aquella que refleje su actitud en mayor medida.

EJEMPLOS DE ESCALAS BIPOLARES

Objeto de actitud: Candidato "A"



Debe observarse que los adjetivos son "extremos" y que entre ellos hay siete opciones de respuesta. Cada sujeto califica al candidato "A" en términos de esta escala de adjetivos bipolares.

⁴² Para profundizar en el diferencial semántico se recomienda consultar *Osgood, Suci y Tannenbaum* (1957), (1976a) y (1976b); así como *Heise* (1976).

FIGURA 9.11

JEMPLOS DE ADJETIVOS BIPOLARES

fuerte-débil	pcceroso-impctente
grande-pequeño	vivo-muerto
bonito-feo	joven-viejo
alto-bajo	rápido-lento
claro-oscuro	gigante-enano
caliente-frío	perfecto-imperfecto
costoso-barato	agradable-desagradable
activo-pasivo	bendito-maldito
seguro-peligroso	útil-inútil
bueno-malo	favorable-desfavorable
dulce-ácido	agresivo-tímido
profundo-superficial	

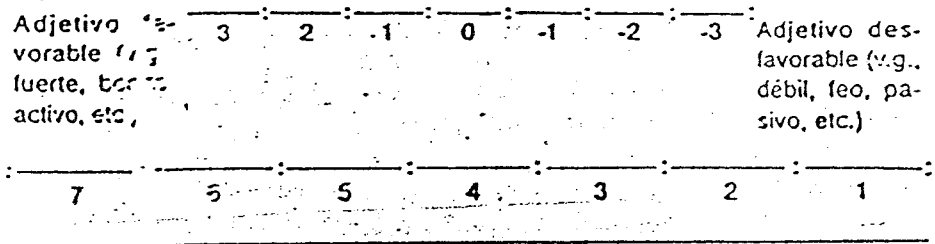
La figura 9.11 presenta sólo algunos ejemplos, desde luego hay muchos más que han sido utilizados o que pudieran pensarse. La elección de adjetivos depende del objeto de actividad a calificar, los adjetivos deben poder aplicarse a éste.

CODIFICACIÓN DE LAS ESCALAS

Los puntos y categorías de la escala pueden codificarse de diversos modos, éstos se presentan en la figura 9.12.

FIGURA 9.12

MANERAS COMUNES DE CODIFICAR EL DIFERENCIAL SEMÁNTICO.



En los casos en que los respondientes tengan menor capacidad de discriminación, se pueden reducir las categorías a cinco opciones. Por ejemplo:

- 5. Seleccionamos los ítems que presenten correlaciones significativas con los demás ítems. Naturalmente, si hay confiabilidad y validez, estas correlaciones serán significativas.
- 6. Desarrollamos la versión final de la escala.

La escala final se califica de igual manera que Likert: sumando las puntuaciones obtenidas respecto a cada ítem o par de adjetivos. La figura 9.14 es un ejemplo de ello.

FIGURA 9.14

EJEMPLO DE CÓMO CALIFICAR UN DIFERENCIAL SEMÁNTICO

sabroso	:	—	:	X	:	—	:	—	:	—	:	—	:	—	:	—	:	insípido
rico	:	X	:	—	:	—	:	—	:	—	:	—	:	—	:	—	:	pobre
suave	:	—	:	X	:	—	:	—	:	—	:	—	:	—	:	—	:	áspero
balanceado	:	—	:	X	:	—	:	—	:	—	:	—	:	—	:	—	:	désbalanceado

valor = 6 + 7 + 6 + 6 = 25

Su interpretación depende del número de ítems o pares de adjetivos. Asimismo, en ocasiones se califica el promedio obtenido en la escala total $\left(\frac{\text{Puntuación total}}{\text{número de ítems}} \right)$. Y se pueden utilizar distintas escalas o diferenciales semánticos para medir actitudes hacia varios objetos. Por ejemplo, podemos medir con cuatro pares de adjetivos la actitud hacia el candidato "A", con otros tres pares de adjetivos la actitud respecto a su plataforma ideológica y con otros seis pares de adjetivos la actitud hacia su partido político. Tenemos tres escalas, cada una con distintos pares de adjetivos para medir la actitud en relación a tres diferentes objetos.

El diferencial semántico es estrictamente una escala de medición ordinal, pero es común que se le trabaje como si fuera de intervalo.

Escalograma de Guttman⁴³

Este método para medir actitudes fue desarrollado por Louis Guttman. Se basa en el principio de que algunos ítems indican en mayor medida la fuerza o intensidad de la actitud. La escala está constituida por afirmaciones, las cuales poseen las mismas características que en el caso de Likert. Pero el escalograma garantiza que la escala mide una dimensión única. Es decir, cada afirmación mide la misma dimensión de la misma variable, a esta propiedad se le conoce como "unidimensionalidad".

Algunos autores consideran que el escalograma más que ser un método de medición de actitudes, es una técnica para determinar si un conjunto de afirmaciones reúnen los requisitos de un tipo particular de escala (v.g., Edwards — 1957 —).

⁴³ Para profundizar en esta escala se sugiere consultar Nie et al. (1975), Black y Champion (1976), Guttman (1976), Lingoes (1976), Dotson y Summers (1976) y Padua (1979).

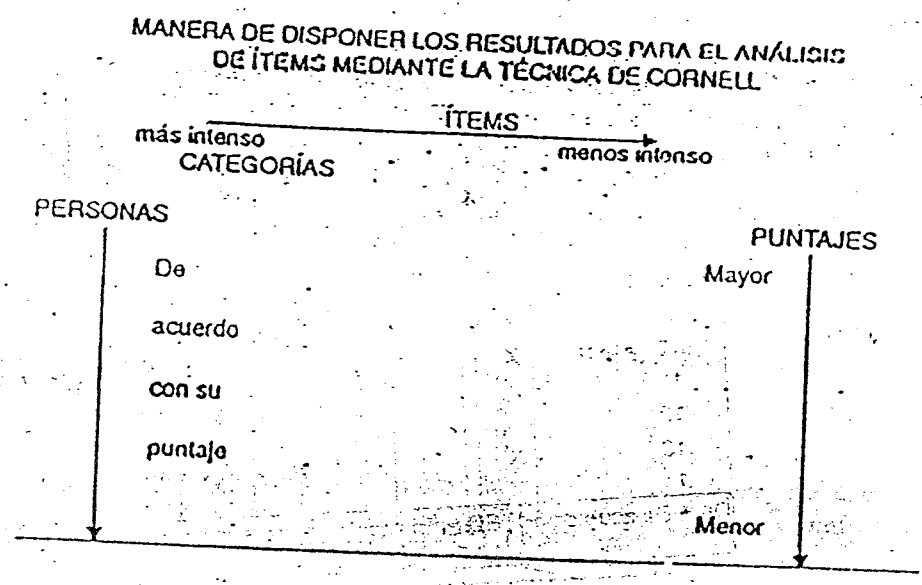
Para construir el escalograma es necesario desarrollar un conjunto de afirmaciones pertinentes al objeto de actitud. Estas deben variar en intensidad. Por ejemplo, si pretendiéramos medir la actitud hacia la calidad en el trabajo dentro del nivel gerencial, la afirmación: "La calidad debe vivirse en todas las actividades del trabajo y en el hogar" es más intensa que la afirmación: "La calidad debe vivirse sólo en las actividades más importantes del trabajo". Dichas afirmaciones se aplican a una muestra a manera de prueba piloto. Y una vez administradas se procede a su análisis. Cabe agregar que las categorías de respuesta para las afirmaciones, pueden variar entre dos ("de acuerdo-en desacuerdo", "si-no", etcétera) o más categorías (v.g., las mismas categorías que en el caso de Likert).

TÉCNICA DE CORNELL

La manera más conocida de analizar los ítems o afirmaciones y desarrollar el escalograma es la técnica Cornell (Guttman, 1976). En ella se procede a:

1. Obtener el puntaje total de cada sujeto en la escala.
2. Ordenar a los sujetos de acuerdo con su puntaje total (del puntaje mayor al menor, de manera vertical descendente).
3. Ordenar a las afirmaciones de acuerdo con su intensidad (de mayor a menor y de izquierda a derecha).
4. Construir una tabla donde se crucen los puntajes de los sujetos ordenados con los ítems y sus categorías jerarquizados (as). Así, tenemos una tabla donde los sujetos constituyen los renglones y las categorías de los ítems forman las columnas. Esto se representa en la figura 9.15.
5. Analizar el número de errores o rupturas en el patrón ideal de intensidad de la escala.

FIGURA 9.15



EJEMPLO

Supongamos que aplicamos una escala con 4 ítems o afirmaciones a 14 sujetos. Cada afirmación tiene dos categorías de respuesta ("de acuerdo" y "en desacuerdo", codificadas como 1 y 0 respectivamente). Los resultados se muestran en la tabla 9.1.

TABLA 9.1

EJEMPLO DE LA TÉCNICA DE CÖRNELL PARA EL ANÁLISIS DE ÍTEMS O AFIRMACIONES

SUJETO	AFIRMACIONES								PUNTUACIONES TOTALES
	A		B		C		D		
	DA (1)	ED (0)	DA (1)	ED (0)	DA (1)	ED (0)	DA (1)	ED (0)	
1	X		X		X		X		4
2	X		X		X		X		4
3	X		X		X		X		4
4		X	X		X		X		3
5		X	X		X		X		3
6		X	X		X		X		3
7		X		X	X		X		2
8		X		X	X		X		2
9		X		X	X		X		2
10		X		X	X		X		2
11		X		X		X	X		2
12		X		X		X	X		1
13		X		X		X		X	1
14		X		X		X		X	0

DA = De acuerdo o 1. ED = En desacuerdo o 0.

Como puede observarse en la tabla 9.1, los sujetos están ordenados por su puntuación en la escala total. Asimismo, las frases deben estar ordenadas por su intensidad (en el ejemplo, A tiene mayor intensidad que B, B mayor que C y C mayor que D) y también sus categorías se encuentran jerarquizadas de acuerdo con su valer de izquierda a derecha. Hay que recordar que si la afirmación es negativa, la codificación se invierte ("La calidad es poco importante para el desarrollo de una empresa", "de acuerdo" se codificaría con cero y "en desacuerdo" con uno). En el ejemplo de la tabla 9.1 tenemos cuatro afirmaciones positivas.

Los sujetos que estén "de acuerdo" con la afirmación "A", que es la más intensa, es muy probable que también estén de acuerdo con las afirmaciones "B", "C" y "D", ya que su intensidad es menor. Los individuos que respondan "de acuerdo" a la

afirmación "B", tenderán a estar "de acuerdo" con "C" y "D" (afirmaciones menos intensas), pero no necesariamente con "A". Quienes estén "de acuerdo" con "C", lo más probable es que se encuentren "de acuerdo" con "D", pero no necesariamente con "A" y "B".

Debe observarse que el sujeto número 1 estuvo "de acuerdo" respecto a las cuatro afirmaciones. Los sujetos 2 y el 3 respondieron de igual forma. Las puntuaciones de todos ellos equivalen a 4 (1+1+1+1). Los sujetos 4, 5 y 6 obtuvieron una puntuación de 3, pues estuvieron "de acuerdo" con tres afirmaciones y así sucesivamente (los últimos dos sujetos estuvieron "en desacuerdo" respecto a todas las afirmaciones). Idealmente, los sujetos que obtienen una puntuación total de 4 en esta particular escala han respondido "de acuerdo" a las cuatro afirmaciones. Los individuos que alcanzan una puntuación total de 3 han respondido estar "en desacuerdo" con la primera afirmación pero están "de acuerdo" con las demás afirmaciones. Quienes reciben una puntuación de 2 manifiestan estar "en desacuerdo" con los dos primeros ítems pero "de acuerdo" con los dos últimos. Los sujetos con puntuación de 1, han respondido "en desacuerdo" a las tres primeras afirmaciones y "de acuerdo" a la última. Finalmente aquellos que hayan estado "en desacuerdo" respecto a las cuatro afirmaciones, tienen una puntuación total de 0.

Los sujetos se escalan de manera perfecta, sin que nadie rompa el patrón de intensidad de las afirmaciones: si están "de acuerdo" con una afirmación más intensa, de igual manera lo estarán con las menos intensas. Por ejemplo, si estoy de acuerdo con la afirmación "Podría casarme con una persona de nivel económico diferente al mío", seguramente estaré de acuerdo con la afirmación "Podría viajar en un automóvil con una persona de nivel económico diferente al mío" (casarse es más intenso que viajar).

Cuando los individuos se escalan perfectamente respecto a las afirmaciones, esto quiere decir que los ítems verdaderamente varían gradualmente en intensidad. Es la prueba empírica de que están escalados por su intensidad. Se le denomina "reproductividad" al grado en que un conjunto de afirmaciones o ítems escalan perfectamente en cuanto a intensidad. Esto significa, que el patrón de respuesta de una persona en relación a todos los ítems puede ser reproducido con exactitud, simplemente conociendo su puntuación total en toda la escala (Black y Champion, 1976).

La reproductividad ideal se da cuando nadie rompe el patrón de intensidad de la escala. Sin embargo en la realidad, sólo unas cuantas escalas del tipo de Guttman reúnen la reproductividad ideal, la mayoría contienen inconsistencias o rupturas al patrón de intensidad. El grado en que se alcanza el patrón perfecto de intensidad de la escala o reproductividad, se determina analizando el número de personas o casos que rompen dicho patrón, que es el quinto paso para construir el escalograma de Guttman.

ANÁLISIS DEL NÚMERO DE ERRORES O RUPTURAS AL PATRÓN IDEAL DE INTENSIDAD DE LA ESCALA

Un error es una inconsistencia en las respuestas de una persona a una escala, es un rompimiento con el patrón ideal de intensidad de ésta. La Tabla 9.2 muestra tres

ejemplos de error (encerrados en círculos), y como puede verse son inconsistencias al patrón ideal. El segundo sujeto respondió "de acuerdo" a los ítems más intensos o fuertes y "en desacuerdo" al ítem menos intenso. El cuarto sujeto manifestó estar "de acuerdo" con las afirmaciones "B" y "C" —supuestamente más intensas— pero "en desacuerdo" con "D" —supuestamente menos intensa—. El quinto sujeto estuvo "en desacuerdo" con los ítems menos intensos pero "de acuerdo" con el más intenso. Son inconsistencias o errores. Si un escalograma presenta diversos errores significa que los ítems no tienen verdaderamente distintos niveles de intensidad.

TABLA 9.2

EJEMPLOS DE ERRORES O INCONSISTENCIAS EN UN ESCALOGRAMA DE GUTTMAN

SUJETO	AFIRMACIONES								PUNTUACIONES TOTALES
	A		B		C		D		
	DA	ED	DA	ED	DA	ED	DA	ED	
1	X		X		X		X		4
2	X		X		X			(X)	3
3		X	X		X		X		3
4		X	X		X			(X)	2
5	(X)			X		X		X	1
6		X		X		X		X	0

Los errores se detectan analizando las respuestas que rompen el patrón y para ello se establecen los "puntos de corte" en la tabla donde se cruzan las afirmaciones y sus categorías con las puntuaciones totales. En el ejemplo de la tabla 9.1 los "puntos de corte" serían los que se muestran en la tabla 9.3.

TABLA 9.3

EJEMPLO DE ESTABLECIMIENTO DE LOS PUNTOS DE CORTE EN LA TÉCNICA DE CORNELL

SUJETO	AFIRMACIONES								PUNTUACIONES TOTALES
	A		B		C		D		
	DA	ED	DA	ED	DA	ED	DA	ED	
1	(1)	(0)	(1)	(0)	(1)	(0)	(1)	(0)	4
	X		X		X		X		

13		X		X		X		X	0
14		X		X		X		X	0

----- puntos de corte (líneas punteadas)

Como se han mencionado anteriormente, cuando el número de errores es excesivo, la escala no presenta *reproductividad* y no puede aceptarse. La *reproductividad* se determina mediante un coeficiente. La fórmula de este coeficiente es:

$$\text{Coeficiente de reproductividad} = \frac{\text{Número de errores o inconsistencias}}{\text{Número total de respuestas}}$$

donde el número total de respuestas = número de ítems o afirmaciones X número de sujetos. Por lo tanto, la fórmula directa sería:

$$\text{Coeficiente de reproductividad} = \frac{\text{Número de errores}}{(\text{número de ítems}) (\text{número de sujetos})}$$

En el ejemplo de la tabla 9.4 tendríamos que el coeficiente de reproductividad es:

$$Cr = 1 - \frac{4}{(4)(14)}$$

$$Cr = 1 - 0.07$$

$$Cr = .93$$

El coeficiente de reproductividad oscila entre 0 y 1, y cuando equivale a .90 o más nos indica que el número de errores es tolerable y la escala es unidimensional y se acepta. Cuando dicho coeficiente es menor a .90 no se acepta la escala. Guttman originalmente recomendó administrar un máximo de 10 a 12 ítems o afirmaciones a un mínimo de 100 personas (Black y Champion, 1976).

Una vez determinado el número de errores aceptable mediante el coeficiente de reproductividad, se procede a aplicar la escala definitiva (si dicho coeficiente fue de .90 o más, esto es, el error permitido no excedió al 10%) o a hacer ajustes en la escala (reconstruir ítems, eliminar ítems que estén generando errores, etcétera). Los cinco pasos mencionados son una especie de prueba piloto para demostrar que la escala es unidimensional y funciona.

CODIFICACIÓN DE RESPUESTAS

Quando se aplica la versión definitiva de la escala los resultados se *codifican* de la misma manera que en la escala Likert, dependiendo del número de categorías de respuesta que se incluyan. Y al igual que la escala Likert y el diferencial semántico, *todas las ítems deben tener el mismo número de categorías de respuesta. Este es un requisito de todas las escalas de actitud.* Asimismo, se considera una respuesta inválida a quien marque dos o más opciones para una misma afirmación. El *escalo-*

grama de Guttman es una escala estrictamente ordinal que puede trabajar como si fuera de intervalo. Puede aplicarse mediante un sistema de tarjetas conteniendo las opciones o categorías de respuesta— o puede ser autoadministrada.

9.6.2. Cuestionarios

Tal vez el instrumento más utilizado para recolectar los datos es el cuestionario. Un cuestionario consiste en un conjunto de preguntas respecto a una o más variables a medir.

¿Qué tipos de preguntas puede haber?

El contenido de las preguntas de un cuestionario puede ser tan variado como los aspectos que se midan a través de éste. Y básicamente, podemos hablar de dos tipos de preguntas: "cerradas" y "abiertas".

Las preguntas "cerradas" contienen categorías o alternativas de respuesta que han sido delimitadas. Es decir, se presentan a los sujetos las posibilidades de respuesta y ellos deben circunscribirse a éstas. Las preguntas "cerradas" pueden ser dicotómicas (dos alternativas de respuesta) o incluir varias alternativas de respuesta. Ejemplos de preguntas cerradas dicotómicas serían: ¿Estudia usted actualmente?

- Sí
- No

¿Durante la semana pasada vio la telenovela "Los Amantes"?

- Sí
- No

Ejemplos de preguntas "cerradas" con varias alternativas de respuesta serían:
¿Cuánta televisión ves los domingos?

- No veo televisión
- Menos de una hora
- 1 o 2 horas
- 3 horas
- 4 horas
- 5 horas o más

¿Cuál es el puesto que ocupa en su empresa?

- Director General / Presidente o Director
- Gerente / Subdirector
- Subgerente / Superintendente
- Coordinador
- Jefe de área

- () Supervisor
- () Empleado
- () Obrero
- () Otro

Si usted tuviera elección, ¿preferiría que su salario fuera de acuerdo con su productividad en el trabajo?

- () Definitivamente sí
- () Probablemente sí
- () No estoy seguro
- () Probablemente no
- () Definitivamente no

Como puede observarse, en las preguntas "cerradas" las categorías de respuesta son definidas a priori por el investigador y se le presentan al respondiente, quien debe elegir la opción que describa más adecuadamente su respuesta. Las escalas de actitudes en forma de pregunta caerían dentro de la categoría de preguntas "cerradas".

Ahora bien, hay preguntas "cerradas", donde el respondiente puede seleccionar más de una opción o categoría de respuesta.

EJEMPLO

Esta familia tiene:

- ¿Radio?
- ¿Televisión?
- ¿Videocasetera?
- ¿Teléfono?
- ¿Automóvil o camioneta?
- Ninguno de los anteriores

Algunos respondientes pudieran marcar una, dos, tres, cuatro o cinco opciones de respuesta. Las categorías no son mutuamente excluyentes. Otro ejemplo sería la siguiente pregunta:

De los siguientes servicios que presta la biblioteca, ¿cuál o cuáles utilizaste el semestre anterior? (Puede señalar más de una opción.)

DE LA SALA DE LECTURA:

- No entré
- A consultar algún libro
- A consultar algún periódico
- A estudiar
- A pasar trabajos a máquina
- A buscar a alguna persona
- Otros, especifica _____

DE LA BIBLIOTECA:

00151

- No entrec
- A consultar algún periódico
- A usar las videocassetteras
- A estudiar
- A hacer trabajos
- A sacar fotocopias
- A leer algún libro
- Otros, especifica _____

DEL MOSTRADOR DE PRÉSTAMOS:

- No fui
- A solicitar algún libro
- A solicitar alguna tesis
- A solicitar algún periódico
- A solicitar diapositivas
- A solicitar máquinas de escribir
- A solicitar equipo audiovisual
- A solicitar asesoría para la localización de material
- Otros, especifica _____

En otras ocasiones, el respondiente tiene que jerarquizar opciones. Por ejemplo: ¿cuál de los siguientes conductores de televisión considera usted el mejor?, ¿cuál en segundo lugar?, ¿cuál en tercer lugar?⁴⁴

- LEM
- BCC
- MME

O bien debe asignar un puntaje a una o diversas cuestiones.

EJEMPLO

¿CUÁNTO LE INTERESA DESARROLLAR? (Indique de 1 a 10 en cada caso según sus intereses)

- _____ Administración de sueldos y compensaciones.
- _____ Salud, seguridad e higiene.
- _____ Administración y negociación de contratos.

⁴⁴ Conductores ficticios.

- _____ Relaciones con sindicatos.
- _____ Habilidades de comunicación ejecutiva.
- _____ Programas y procesos sobre calidad/productividad.
- _____ Calidad de vida en el trabajo.
- _____ Teoría de la organización.
- _____ Administración financiera.
- _____ Desarrollo organizacional / innovación.
- _____ Técnicas de investigación organizacional.
- _____ Estructura organizacional (tamaño, complejidad, formalización).
- _____ Sistemas de información y control.
- _____ Auditoría administrativa.
- _____ Planeación estratégica.
- _____ Sistemas de computación.
- _____ Mercadotecnia y comercialización.
- _____ Otros (especificar):
- _____
- _____
- _____

En cambio, las preguntas "abiertas" no delimitan de antemano las alternativas de respuesta: Por lo cual el número de categorías de respuesta es muy elevado. En teoría es infinito.

EJEMPLO

¿Por qué asiste a psicoterapia?

¿Qué opina del programa de televisión "Los Cazadores"?

¿De qué manera la directiva de la empresa ha logrado la cooperación del sindicato para el proyecto de calidad?

¿Usamos preguntas cerradas o abiertas?

Cada cuestionario obedece a diferentes necesidades y problemas de investigación, lo que origina que en cada caso el tipo de preguntas a utilizar sea diferente. Algunas veces se incluyen solamente preguntas "cerradas", otras veces únicamente preguntas

"abiertas" y en ciertos casos ambos tipos de preguntas. *Cada clase de pregunta tiene sus ventajas y desventajas. Las cuales se mencionan a continuación.*

Las preguntas "cerradas" son fáciles de codificar y preparar para su análisis. Asimismo, estas preguntas requieren de un menor esfuerzo por parte de los respondientes. Estos no tienen que escribir o verbalizar pensamientos, sino simplemente seleccionar la alternativa que describa mejor su respuesta. Responder a un cuestionario con preguntas cerradas toma menos tiempo que contestar a uno con preguntas abiertas. Si el cuestionario es enviado por correo, se tiene una mayor respuesta cuando es fácil de contestar y requiere menos tiempo completarlo. La principal desventaja de las preguntas "cerradas" reside en que limitan las respuestas de la muestra y —en ocasiones— ninguna de las categorías describe con exactitud lo que los sujetos tienen en mente, no siempre se captura lo que pasa por las cabezas de los sujetos.

Para poder formular preguntas "cerradas" es necesario anticipar las posibles alternativas de respuesta. De no ser así es muy difícil plantearlas. Asimismo, el investigador tiene que asegurarse que los sujetos a los cuales se les administrarán, conocen y comprenden las categorías de respuesta. Por ejemplo, si preguntamos qué canal de televisión es el preferido, determinar las opciones de respuesta y que los respondientes las comprendan es muy sencillo. Pero si preguntamos sobre las razones y motivos que provocan esa preferencia, determinar dichas opciones es algo bastante más complejo.

Las preguntas "abiertas" son particularmente útiles cuando no tenemos información sobre las posibles respuestas de las personas o cuando esta información es insuficiente. También sirven en situaciones donde se desea profundizar una opinión o los motivos de un comportamiento. Su mayor desventaja es que son más difíciles de codificar, clasificar y preparar para su análisis. Además, pueden presentarse sesgos derivados de distintas fuentes: por ejemplo, quienes tienen dificultades para expresarse oralmente y por escrito pueden no responder con precisión lo que realmente desean o generar confusión en sus respuestas. El nivel educativo, la capacidad de manejo del lenguaje y otros factores pueden afectar la calidad de las respuestas (Black y Champion, 1976). Asimismo, responder a preguntas "abiertas" requiere de un mayor esfuerzo y tiempo.

La elección del tipo de preguntas que contenga el cuestionario depende del grado en que se puedan anticipar las posibles respuestas, los tiempos de que se disponga para codificar y si se quiere una respuesta más precisa o profundizar en alguna cuestión. Una recomendación para construir un cuestionario es que se analice variable por variable qué tipo de pregunta o preguntas pueden ser más confiables y válidas para medir a esa variable, de acuerdo con la situación del estudio (planteamiento del problema, características de la muestra, análisis que se piensan efectuar, etcétera).

¿Una o varias preguntas para medir una variable?

En ocasiones sólo basta una pregunta para recolectar la información necesaria sobre la variable a medir. Por ejemplo, para medir el nivel de escolaridad de una muestra,

basta con preguntar: ¿Hasta qué año escolar cursó? o ¿cuál es su grado máximo de estudios? En otras ocasiones es necesario elaborar varias preguntas para verificar la consistencia de las respuestas. Por ejemplo, el nivel económico puede medirse preguntando: Aproximadamente ¿cuál es su nivel mensual de ingresos? y preguntando: Aproximadamente, ¿cuántos focos eléctricos tiene en su casa?⁴³ Además de preguntar sobre propiedades, inversiones, puesto que ocupa la fuente principal de ingresos de la familia (generalmente, el padre), etcétera.

Al respecto, es recomendable hacer solamente las preguntas necesarias para obtener la información deseada o medir la variable. Si una pregunta es suficiente no es necesario incluir más, no tiene sentido. Si se justifica hacer varias preguntas, entonces es conveniente plantearlas en el cuestionario. Esto último ocurre con frecuencia en el caso de variables con varias dimensiones o componentes a medir, en donde se incluyen varias preguntas para medir las distintas dimensiones. Se tienen varios indicadores.

EJEMPLO

La empresa Comunicometría, S. C., realizó una investigación para la Fundación Mexicana para la Calidad Total, A.C. (1988), con el propósito de conocer las prácticas, técnicas, estructuras, procesos y temáticas existentes en materia de Calidad Total en México. El estudio fue de carácter exploratorio y constituyó el primer esfuerzo por obtener una radiografía del estado de los procesos de calidad en dicho país.

En esta investigación se elaboró un cuestionario que medía el grado en que las organizaciones mexicanas aplicaban diversas prácticas tendientes a elevar la calidad, la productividad y la calidad de vida en el trabajo. Una de las variables importantes era el "grado en que se distribuía la información sobre el proceso de calidad en la organización". Esta variable se midió a través de las siguientes preguntas:

- A. Por lo que respecta a los programas de información sobre calidad, ¿cuáles de las siguientes actividades se efectúan en esta empresa?
- (1) Planeación del manejo de datos sobre calidad.
 - (2) Formas de control.
 - (3) Elaboración de reportes con datos sobre calidad.
 - (4) Evaluación sistemática de los datos sobre calidad.
 - (5) Distribución generalizada de información sobre calidad.
 - (6) Sistemas de autocontrol de calidad.
 - (7) Distribución selectiva de datos sobre calidad.

- B. Sólo a quienes distribuyen selectivamente datos sobre calidad: ¿A qué niveles de la empresa?
-
-
-

⁴³ En varios estudios se ha demostrado que el nivel de ingresos está relacionado con el número de focos de una casa/habitación. El número de focos está vinculado con el número de cuartos de la casa, extensión de ésta, presencia de focos en el jardín de la casa, candeleros y otros factores.

C. ¿Sólo a quienes distribuyen selectivamente datos sobre calidad: ¿a que funciones? _____

D. ¿Qué otras actividades se realizan en esta empresa para los programas de información sobre calidad? _____

En este ejemplo, las preguntas "B" y "C" se elaboraron para ahondar sobre los receptores o usuarios de los datos en aspectos del control de calidad distribuidos selectivamente. Se justifica el hacer estas dos preguntas, ayuda a tener mayor información sobre la variable. Cuando se tienen varias preguntas para una misma variable se dice que se tiene una "batería de preguntas".

¿Las preguntas van precodificadas o no?

Siempre que se pretendan efectuar análisis estadísticos es necesario codificar las respuestas de los sujetos a las preguntas del cuestionario, y debemos recordar que esto significa asignarles símbolos o valores numéricos a dichas respuestas. Ahora bien, cuando se tienen preguntas "cerradas", es posible codificar "a priori" o precodificar las alternativas de respuesta e incluir esta precodificación en el cuestionario (tal y como lo hacemos con las escalas de actitudes).

EJEMPLOS DE PREGUNTAS PRECODIFICADAS

¿Tiene usted Inversiones en la Bolsa de Valores?

1 Si

0 No

Quando se enfrenta usted a un problema en su trabajo, para resolverlo recurre generalmente a:

- (1) Su superior inmediato
- (2) Su propia experiencia
- (3) Sus compañeros
- (4) Los manuales de políticas y procedimientos
- (5) Otra fuente _____

(especificar)

En ambas preguntas, las respuestas van acompañadas de su valor numérico correspondiente, han sido precodificadas. Obviamente en las preguntas "abiertas" no puede darse la precodificación, la codificación se realiza posteriormente, una vez que se tienen las respuestas. Las preguntas y alternativas de respuesta precodificadas tienen la ventaja que su codificación y preparación para el análisis son más sencillas y requieren de menos tiempo.

¿Qué características debe tener una pregunta?

Independientemente de que las preguntas sean abiertas o cerradas y sus respuestas estén precodificadas o no, hay una serie de características que deben cubrirse al plantearlas:

- A. *Las preguntas deben ser claras y comprensibles para los respondientes.* Deben evitarse términos confusos o ambiguos y como menciona Rojas (1981, p.138) no es nada recomendable sacrificar la claridad por concisión. Es indispensable incluir las palabras que sean necesarias para que se comprenda la pregunta. Desde luego, sin ser repetitivos o barrocos. Por ejemplo, la pregunta: ¿ve usted televisión? es confusa, no delimita cada cuánto. Sería mucho mejor especificar: ¿acostumbra usted ver televisión diariamente? o ¿cuántos días durante la última semana vio televisión? y después preguntar los horarios, canales y contenidos de los programas.
- B. *Las preguntas no deben incomodar al respondiente.* Preguntas como: ¿acostumbra consumir algún tipo de bebida alcohólica?, tienden a provocar rechazo. Es mejor preguntar: ¿algunos de sus amigos acostumbra consumir algún tipo de bebida alcohólica? y después utilizar preguntas sutiles que indirectamente nos indiquen si la persona acostumbra consumir bebidas alcohólicas (v.g., ¿cuál es su tipo de bebida favorita?, etcétera). Y hay temáticas en donde a pesar de que se utilicen preguntas sutiles, el respondiente se sentirá molesto. En estos casos, pueden utilizarse escalas de actitud en lugar de preguntas o aún otras formas de medición. Tal es el caso de temas como el homosexualismo, la prostitución, la pornografía, los anticonceptivos y la drogadicción.
- C. *Las preguntas deben preferentemente referirse a un sólo aspecto o relación lógica.* Por ejemplo, la pregunta: ¿acostumbra usted ver televisión y escuchar radio diariamente?, expresa dos aspectos, puede confundir. Es mucho mejor dividirla en dos preguntas, una relacionada con la televisión y otra relacionada con la radio.
- D. *Las preguntas no deben inducir las respuestas* (Rojas, 1981, p. 138). Preguntas tendenciosas o que dan pie a elegir un tipo de respuesta deben evitarse. Por ejemplo: ¿considera usted a Ricardo Hernández el mejor candidato para dirigir nuestro sindicato?, es una pregunta tendenciosa, induce la respuesta. Lo mismo que la pregunta: ¿los trabajadores mexicanos son muy productivos? Se insinúa la respuesta en la pregunta. Resultaría mucho más conveniente preguntar:

¿Qué tan productivos considera usted —en general— a los trabajadores mexicanos?

Sumamente productivos

Más bien productivos

Más bien improductivos

Sumamente improductivos

- E. *Las preguntas no pueden apoyarse en instituciones, ideas respaldadas socialmente ni en evidencia comprobada.* Es también una manera de inducir respuestas. Por ejemplo, la pregunta: ¿la Organización Mundial de la Salud ha realizado

diversos estudios y concluyó que el tabaquismo provoca diversos daños al organismo. ¿usted considera que fumar es nocivo para su salud? Esquemas del tipo: "La mayoría de las personas opinan que...", "La Iglesia considera...", "Los padres de familia piensan que...", etcétera, no deben anteceder a una pregunta, sesgan las respuestas.

- F. En las preguntas con varias alternativas o categorías de respuesta y donde el respondiente sólo tiene que elegir una, puede ocurrir que el orden en que se presenten dichas alternativas afecte las respuestas de los sujetos (v.g., tiendan a favorecer a la primera o a la última alternativa de respuesta). Entonces resulta conveniente rotar el orden de lectura de las alternativas de manera proporcional. Por ejemplo, si preguntamos: ¿cuál de los siguientes tres candidatos presidenciales considera usted que logrará disminuir verdaderamente la inflación? Y el 33.33% de las veces que se haga la pregunta se menciona primero al candidato "A", el 33.33% se menciona primero al candidato "B" y el restante 33.33% al candidato "C".
- G. El lenguaje utilizado en las preguntas debe ser adaptado a las características del respondiente (tomar en cuenta su nivel educativo, socioeconómico, palabras que maneja, etcétera). Este aspecto es igual al que se comentó sobre las escalas de actitudes.

¿Cómo deben ser las primeras preguntas de un cuestionario?

En algunos casos es conveniente iniciar con preguntas neutrales o fáciles de contestar, para que el respondiente vaya adentrándose en la situación. *No se recomienda comenzar con preguntas difíciles de responder o preguntas muy directas.* Imaginemos un cuestionario diseñado para obtener opiniones en torno al aborto que empiece con una pregunta poco sutil tal como: ¿Está usted de acuerdo en que se legalice el aborto en este país? Sin lugar a dudas será un fracaso.

A veces los cuestionarios pueden comenzar con preguntas demográficas sobre el estado civil, sexo, edad, ocupación, nivel de ingresos, nivel educativo, religión, ideología, puesto en una organización o algún tipo de afiliación a un grupo, partido o institución. Pero en otras ocasiones es mejor hacer este tipo de preguntas al final del cuestionario, particularmente cuando los sujetos puedan sentir que se comprometen al responder al cuestionario. Cuando construimos un cuestionario es indispensable que pensemos en cuáles son las preguntas ideales para iniciar. Éstas deberán lograr que el respondiente se concentre en el cuestionario.

¿De qué está formado un cuestionario?

Además de las *preguntas y categorías de respuestas*, un cuestionario está formado por *instrucciones* que nos indican cómo contestar, por ejemplo:

Hablando de la mayoría de sus proveedores en qué medida conoce usted (MOSTRAR TARJETA UNO Y MARCAR LA RESPUESTA EN CADA CASO):

	Comp. (5)	Bast. (4)	Reg. (3)	Poco (2)	Nada (1)
• ¿Las políticas de su proveedor?					
• ¿Sus finanzas (estado financiero)?					
• ¿Los objetivos de su área de ventas?					
• ¿Sus programas de capacitación para vendedores?					
• ¿Número de empleados de su área de ventas?					
• ¿Problemas laborales?					
• ¿Los métodos de producción que tienen?					
• ¿Otros clientes de ellos?					
• ¿Su índice de rotación personal?					

¿Tiene este ejido o comunidad, ganado, aves o colmenas que sean de propiedad colectiva? (CIRCULE LA RESPUESTA)

Sí 1
(continúa)

No 2
(pase a 30)

¿Se ha obtenido la cooperación de todo el personal o la mayoría de éste para el proyecto de calidad?

1 Sí
(pase a la pregunta 26)

2 No
(pase a la pregunta 27)

Las instrucciones son tan importantes como las preguntas y es necesario que sean claras para los usuarios a quienes van dirigidas. Y una instrucción muy importante es agradecer al respondiente por haberse tomado el tiempo de contestar el cuestionario. También, es frecuente incluir una carátula de presentación o una carta donde se expliquen los propósitos del cuestionario y se garantice la confidencialidad de la información, esto ayuda a ganar la confianza del respondiente. A continuación en la figura 9.16 se presentan algunos textos ilustrativos de cartas introductorias a un cuestionario.

EJEMPLOS DE CARTAS

Buenos días (tardes):

Estamos trabajando en un estudio que servirá para elaborar una tesis profesional acerca de la Biblioteca de la Universidad Anáhuac.

Quisiéramos pedir tu ayuda para que contestes a unas preguntas que no llevarán mucho tiempo. Tus respuestas serán confidenciales y anónimas.

Las personas que fueron seleccionadas para el estudio no se eligieron por su nombre sino al azar.

Las opiniones de todos los encuestados serán sumadas y reportadas en la tesis profesional, pero nunca se reportarán datos individuales.

Te pedimos que contestes este cuestionario con la mayor sinceridad posible. No hay respuestas correctas ni incorrectas.

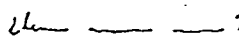
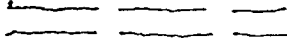
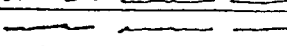
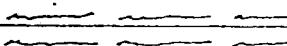
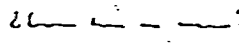
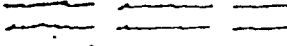
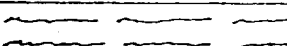
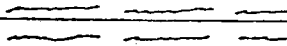
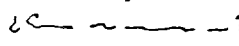
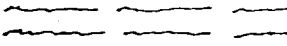
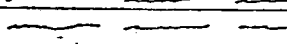
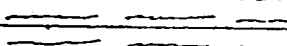
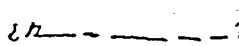
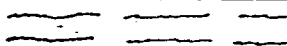
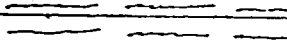
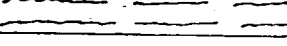

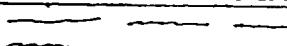
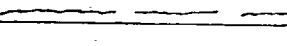
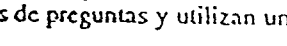
Lee las instrucciones cuidadosamente, ya que existen preguntas en las que sólo pueden responder a una opción; otras son de varias opciones y también se incluyen preguntas abiertas.

Muchas gracias por tu colaboración.


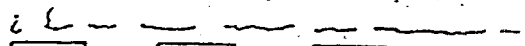
BUENOS DÍAS (TARDES)

COMUNICOMETRÍA ESTÁ HACIENDO UNA ENCUESTA CON EL PROPÓSITO DE CONOCER UNA SERIE DE OPINIONES QUE SE TIENEN ACERCA DE ESTA EMPRESA, Y PARA ELLO LE PEDIRÍA FUERA TAN AMABLE DE CONTESTAR UNAS PREGUNTAS. NO LE TOMARÁ MÁS DE 20 MINUTOS. LA INFORMACIÓN QUE NOS PROPORCIONE SERÁ MANEJADA CON LA MÁS ESTRICTA CONFIDENCIALIDAD. DESDE LUEGO, NO HAY PREGUNTAS DELICADAS.

La manera en que pueden distribuirse las preguntas, categorías de respuestas o instrucciones es variada. Algunos prefieren colocar las preguntas a la izquierda y las respuestas a la derecha, con lo que se tendría un formato como el siguiente:

¿U...? 	<input type="checkbox"/>			
¿L...? 	<input type="checkbox"/>			
¿C...? 	<input type="checkbox"/>			
¿H...? 	<input type="checkbox"/>			
¿S...? 	<input type="checkbox"/>			

Otros dividen el cuestionario por secciones de preguntas y utilizan un formato horizontal. Por ejemplo:

PRESENTACIÓN			
PREGUNTAS SOBRE MOTIVACIÓN INTRÍNSECA:			
¿E...? 	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
¿L...? 	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PREGUNTAS SOBRE SATISFACCIÓN LABORAL			
ETC.			

O combinan diversas posibilidades, distribuyendo preguntas que miden la misma variable a través de todo el cuestionario. Cada quien puede utilizar el formato que desee o juzgue conveniente, lo importante es que sea totalmente comprensible para el usuario; las instrucciones, preguntas y respuestas se diferencien, no resulte visualmente tedioso y se pueda leer sin dificultades.

¿De qué tamaño debe ser un cuestionario?

Na existe una regla al respecto, aunque como menciona Padua (1979), si es muy corto se pierde información y si resulta largo puede resultar tedioso de responder. En este último caso, las personas pueden negarse a responder o —al menos— no completar el cuestionario. El tamaño depende del número de variables y dimensiones a medir, el interés de los respondientes y la manera como es administrado (de este punto se hablará en el siguiente inciso). Cuestionarios que duran más de 35 minutos pueden resultar tediosos a menos que los respondientes estén muy motivados para contestar

(v.g., cuestionarios de personalidad, cuestionarios para obtener un trabajo). Una recomendación que puede ayudarnos para evitar un cuestionario más largo de lo requerido es: "No hace preguntas innecesarias o injustificadas".

¿Cómo se codifican las preguntas abiertas?

Las preguntas abiertas se codifican una vez que conocemos todas las respuestas de los sujetos a las cuales se les aplicaron o al menos las principales tendencias de respuestas en una muestra de los cuestionarios aplicados.

El procedimiento consiste en encontrar y darle nombre a los patrones generales de respuesta (respuestas similares o comunes), listar estos patrones y después asignar un valor numérico o símbolo a cada patrón. Así, un patrón constituirá una categoría de respuesta. Para cerrar las preguntas abiertas, se sugiere el siguiente procedimiento, basado parcialmente en Rojas (1981, pp.150-151):

1. Seleccionar determinado número de cuestionarios mediante un método adecuado de muestreo, asegurando la representatividad de los sujetos investigados.
2. Observar la frecuencia con que aparece cada respuesta a la pregunta.
3. Elegir las respuestas que se presentan con mayor frecuencia (patrones generales de respuesta).
4. Clasificar las respuestas elegidas en temas, aspectos o rubros, de acuerdo con un criterio lógico, cuidando que sean mutuamente excluyentes.
5. Darle un nombre o título a cada tema, aspecto o rubro (patrón general de respuesta).
6. Asignarle el código a cada patrón general de respuesta.

Por ejemplo, en la investigación de Comunicometría (1988) se hizo una pregunta abierta:

¿De qué manera la alta gerencia busca obtener la cooperación del personal para el desarrollo del proyecto de calidad?

Las respuestas fueron múltiples pero pudieron encontrarse los siguientes patrones generales de respuesta:

CÓDIGOS	CATEGORÍAS (PATRONES O RESPUESTAS CON MAYOR FRECUENCIA DE MENCIÓN)	NÚM. DE FRECUENCIAS DE MENCIÓN
1	Involucrando al personal y comunicándose con él.	28
2	Motivación e Integración.	20

3	Capacitación en general.	12
4	Incentivos / recompensas.	11
5	Difundiendo el valor "calidad" o la filosofía de la empresa.	7
6	Grupos o sesiones de trabajo.	5
7	Posicionamiento del área de calidad o equivalente.	3
8	Sensibilización.	2
9	Desarrollo de la calidad de vida en el trabajo.	2
10	Incluir aspectos de calidad en el manual de inducción.	2
11	Enfatizar el cuidado de la maquinaria.	2
12	Trabajando bajo un buen clima laboral.	2
13	Capacitación "en cascada".	2
14	Otras.	24

Como varias categorías o patrones tenían solamente dos frecuencias, éstas a su vez pudieron reducirse a:

CATEGORÍAS
Involucrando al personal y comunicándose con él.
Motivación e integración / mejoramiento del clima laboral.
Capacitación.
Incentivos / recompensas.
Difundiendo el valor "calidad" o la filosofía de la empresa.
Grupos o sesiones de trabajo.
Otras.

Al "cerrar" preguntas abiertas y ser codificadas, debe tenerse en cuenta que un mismo patrón de respuesta puede expresarse con diferentes palabras. Por ejemplo, ante la pregunta: ¿Qué sugerencias podría hacer para mejorar al programa "Estelar"? Las respuestas: "Mejorar las canciones y la música", "cambiar las canciones", "incluir nuevas y mejores canciones", etc., pueden agruparse en la categoría o patrón de respuesta "modificar la musicalización del programa".

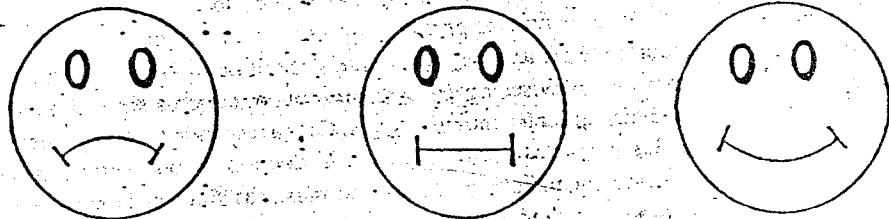
¿En qué contextos puede administrarse o aplicarse un cuestionario?

Los cuestionarios pueden ser aplicados de diversas maneras:

- A) *Autoadministrado*. En este caso el cuestionario se les proporciona directamente a los respondientes, quienes lo contestan. No hay intermediarios y las respuestas las marcan ellos. Por ejemplo, si los respondientes fueran una muestra de los estudiantes de la Licenciatura en Comunicación de Bogotá, se acudiría a ellos y se les entregarían los cuestionarios. Los estudiantes se autoadministrarían el cuestionario. Obviamente que esta manera de aplicar el cuestionario es impropia para analfabetas, personas que tienen dificultades de lectura o niños que todavía no leen adecuadamente.
- B) *Por entrevista personal*. En esta situación, un entrevistador aplica el cuestionario a los respondientes (entrevistados). El entrevistador va haciéndole las preguntas al respondiente y va anotando las respuestas. Las instrucciones son para el entrevistador. Normalmente se tienen varios entrevistadores, quienes deberán estar capacitados en el arte de entrevistar y conocer a fondo el cuestionario, y no deben sesgar o influir las respuestas.
- C) *Por entrevista telefónica*. Esta situación es similar a la anterior, solamente que la entrevista no es "cara a cara" sino a través del teléfono. El entrevistador le hace las preguntas al respondiente por este medio de comunicación.
- D) *Autoadministrado y enviado por correo postal, electrónico o servicio de mensajería*. En este caso también los respondientes contestan directamente el cuestionario, ellos marcan o anotan las respuestas, no hay intermediario. Solamente que no se entregan los cuestionarios directamente a los respondientes ("en propia mano") sino que se les envían por correo u otro medio, no hay retroalimentación inmediata, si los sujetos tienen alguna duda no se les puede aclarar en el momento.

Consejos para la administración del cuestionario, dependiendo del contexto

Cuando se tiene población analfabeta, con niveles educativos bajos o niños que apenas comienzan a leer o no dominan la lectura, el método más conveniente de administración de un cuestionario es por entrevista. Aunque hoy en día ya existen algunos cuestionarios muy gráficos que usan escalas sencillas. Como por ejemplo:



Con trabajadores de niveles de lectura básica se recomienda utilizar entrevistas o cuestionarios autoadministrados sencillos que se apliquen en grupos con la asesoría de entrevistadores o supervisores capacitados.

En algunos casos, con ejecutivos que difícilmente pueden dedicarle a un solo asunto más de 20 minutos, se pueden utilizar cuestionarios autoadministrados o entrevistas telefónicas. Con estudiantes suelen funcionar los cuestionarios autoadministrados.

Asimismo, algunas asociaciones hacen encuestas por correo y ciertas empresas envían cuestionarios a sus ejecutivos y supervisores mediante el servicio interno de mensajería o por correo electrónico. Cuando el cuestionario contiene unas cuantas preguntas (su administración no toma más de 4 minutos —o máximo 5—), la entrevista telefónica es una buena alternativa.

Ahora bien, sea cual fuere la forma de administración, siempre debe haber uno o varios supervisores que verifiquen que se están aplicando correctamente los cuestionarios. Cuando un cuestionario o escala es aplicado(a) de forma masiva suele denominarse "encuesta".

Cuando se utiliza la entrevista telefónica se debe tomar en cuenta el horario. Ya que si hablamos sólo a una hora (digamos en la mañana), nos encontraremos con unos cuantos subgrupos de la población (v.g., amas de casa).

Cuando lo enviamos por correo o es autoadministrado directamente, las instrucciones deben pecar de precisas, claras y completas. Y debemos dar instrucciones que motiven al respondiente para que continúe contestando el cuestionario (v.g., ya nada más unas cuantas preguntas, finalmente).

Asimismo, cabe señalar que cuando se trata de entrevista personal, el lugar donde se realice es importante (oficina, casa-habitación, en la calle, etc.). Por ejemplo, Jaffe, Pasternak y Grifel (1983) realizaron un estudio para comparar —entre otros aspectos— las respuestas obtenidas en dos puntos diferentes: en el hogar y en puntos de venta. El estudio se interesaba en la conducta del comprador y los resultados concluyeron que se pueden obtener datos exactos en ambos puntos, pero la entrevista en los puntos de compra-venta es menos costosa.

Las entrevistas personales requieren de una atmósfera apropiada. El entrevistador debe ser amable y tiene que generar confianza en el entrevistado. Cuando se trata de entrevistados del sexo masculino, mujeres simpáticas y agradables suelen resultar excelentes entrevistadores. Quien responde a una entrevista debe concentrarse en las preguntas y estar relajado. Y después de una entrevista debe prepararse un informe que indique si el sujeto se mostraba sincero, la manera como respondió, el tiempo que duró la entrevista, el lugar donde se realizó, las características del entrevistado, los contratiempos que se presentaron y la manera como se desarrolló la entrevista, así como otros aspectos que se consideren relevantes.

La elección del contexto para administrar el cuestionario deberá ser muy cuidadosa y dependerá del presupuesto de que se disponga, el tiempo de entrega de los resultados, los objetivos de la investigación y el tipo de respondientes (edad, nivel educativo, etc.).

Estas maneras de aplicar un cuestionario pueden hacerse extensivas a las escalas de actitudes, sólo que es mucho más difícil en el caso de que se administren por teléfono (pocas frases y alternativas claras de respuesta).

BIBLIOTECA NACIONAL DE MEXICO

00165

Cuando los cuestionarios son muy complejos de contestar o aplicar, suele utilizarse un manual que explica a fondo las instrucciones y cómo debe responderse o ser administrado.

¿Cuál es el proceso para construir un cuestionario?

Siguiendo los pasos para construir un instrumento de medición tendríamos la figura 9.17.

Un aspecto muy importante que es necesario mencionar, reside en que cuando se construye un cuestionario —al igual que otros instrumentos de medición— se debe ser consistente en todos los aspectos. Por ejemplo, si se decide que las instrucciones vayan en mayúsculas o algún tipo de letra especial, todas las instrucciones deberán ser así. Si se prefiere que los códigos de las categorías de respuesta van en recuadro, todas deberán ajustarse a esto. Si no se es consistente, algunos respondientes o entrevistadores pueden desconcertarse.

9.6.3. Análisis del contenido

¿QUÉ ES Y PARA QUÉ SIRVE EL ANÁLISIS DE CONTENIDO?

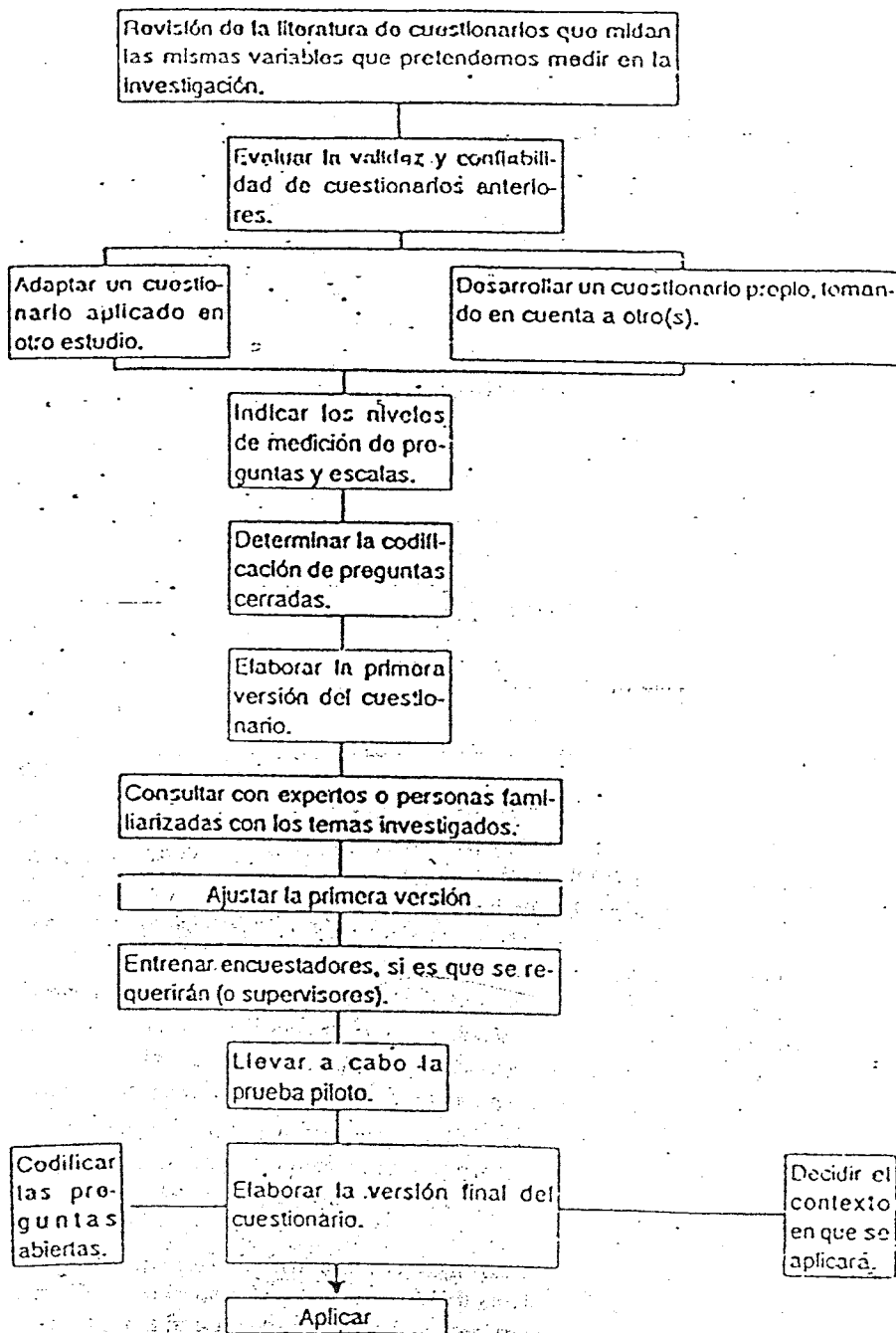
De acuerdo con la definición clásica de Berelson (1952), *el análisis de contenido es una técnica para estudiar y analizar la comunicación de una manera objetiva, sistemática y cuantitativa*. Krippendorff (1982) extiende la definición del análisis de contenido a una *técnica de investigación para hacer inferencias válidas y confiables de datos con respecto a su contexto*. Algunos autores consideran al análisis de contenido como un diseño. Pero más allá de como lo definamos, es una técnica muy útil para analizar los procesos de comunicación en muy diversos contextos. El análisis de contenido puede ser aplicado virtualmente a cualquier forma de comunicación (programas televisivos o radiofónicos, artículos en prensa, libros, poemas, conversaciones, pinturas, discursos, cartas, melodías, reglamentos, etcétera). Por ejemplo, puede servir para analizar la personalidad de alguien, evaluando sus escritos; conocer las actitudes de un grupo de personas mediante el análisis de sus discursos; indagar sobre las preocupaciones de un pintor o un músico; compenetrarse con los valores de una cultura; o averiguar las intenciones de un publicista o propagandista.

USOS DEL ANÁLISIS DE CONTENIDO

Berelson (1952) señala varios usos del análisis de contenido, entre los que destacan:

- 1) *Describir tendencias en el contenido de la comunicación.*
- 2) *Develar diferencias en el contenido de la comunicación (entre personas, grupos, instituciones, países).*
- 3) *Comparar mensajes, niveles y medios de comunicación.*

FIGURA 9.17
PROCESO PARA CONSTRUIR UN CUESTIONARIO



- 4) Auditar el contenido de la comunicación y compararlo contra estándares u objetivos.
- 5) Construir y aplicar estándares de comunicación (políticas, normas, etc.).
- 6) Exponer técnicas publicitarias y de propaganda.
- 7) Medir la claridad de mensajes.
- 8) Descubrir estilos de comunicación.
- 9) Identificar intenciones, apelaciones y características de comunicadores.
- 10) Descifrar mensajes ocultos y otras aplicaciones a la inteligencia militar y a la seguridad política.
- 11) Revelar "centros" de interés y atención para una persona, un grupo y una comunidad.
- 12) Determinar el estado psicológico de personas o grupos.
- 13) Obtener indicios del desarrollo verbal (v.g., en la escuela, como resultado de la capacitación, el aprendizaje de conceptos).
- 14) Anticipar respuestas a comunicaciones.
- 15) Reflejar actitudes, valores y creencias de personas, grupos o comunidades.
- 16) Cerrar preguntas abiertas.

El análisis de contenido puede utilizarse para ver si varias telenovelas difieren entre sí en cuanto a su carga de contenido sexual, para conocer las diferencias ideológicas entre varios periódicos (en términos generales o en referencia a un tema en particular), para comparar estrategias propagandísticas de partidos políticos contendientes, para contrastar a través de sus escritos a diferentes grupos que asisten a psicoterapia, para comparar el vocabulario aprendido por niños que se exponen a mayor contenido televisivo en relación con niños que ven menos televisión, para analizar la evolución de las estrategias publicitarias a través de algún medio respecto a un producto (v.g., perfumes femeninos de costo elevado); para conocer y comparar la posición de diversos presidentes latinoamericanos respecto al problema de la deuda externa; para comparar estilos de escritores que se señalan como parte de una misma corriente literaria; etcétera.

¿CÓMO SE REALIZA EL ANÁLISIS DE CONTENIDO?

El análisis de contenido se efectúa por medio de la codificación, que es el proceso a través del cual las características relevantes del contenido de un mensaje son transformadas a unidades que permitan su descripción y análisis preciso. Lo importante del mensaje se convierte en algo susceptible de describir y analizar. Para poder codificar es necesario definir el universo a analizar, las unidades de análisis y las categorías de análisis.

UNIVERSO

El universo podría ser la obra completa de Franz Kafka, las emisiones de un noticiero televisivo durante un mes, los editoriales publicados en un día por cinco periódicos

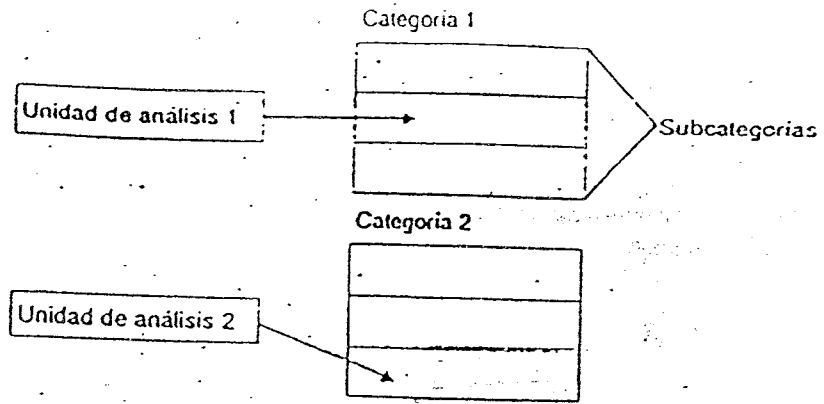
de una determinada ciudad, todos los capítulos de tres telenovelas, los discos de Janis Joplin, Jimi Hendrix y Bob Dylan, los escritos de un grupo de estudiantes durante un ciclo escolar, los discursos pronunciados por varios contendientes políticos durante el último mes previo a la elección, escritos de un grupo de pacientes en psicoterapia, las conversaciones grabadas de 10 parejas que participan en un experimento sobre interacción matrimonial. El universo, como en cualquier investigación, debe delimitarse con precisión.

UNIDADES DE ANÁLISIS

Las unidades de análisis constituyen segmentos del contenido de los mensajes que son caracterizados para ubicarlos dentro de las categorías. Berelson (1952) menciona cinco unidades importantes de análisis:

- 1) *La palabra.* Es la unidad de análisis más simple, aunque como señala Kerlinger (1975), puede haber unidades más pequeñas como letras, fonemas o símbolos. Así, se puede medir cuántas veces aparece una palabra en un mensaje (v.g., veces que en un programa televisivo de fin de año se menciona al Presidente).
- 2) *El tema.* Éste se define a menudo como una oración, un enunciado respecto a algo. Los temas pueden ser más o menos generales. Kerlinger (1975, p. 552) utiliza un excelente ejemplo para ello: "Las cartas de adolescentes o estudiantes de colegios superiores pueden ser estudiadas en sus expresiones de autoreferencia. Éste sería el tema más extenso. Los temas que constituyen éste podrían definirse como cualesquiera oraciones que usen "yo", "mi" y otros términos que indiquen referencia al yo del escritor. Así, se analizaría qué tanta autoreferencia está presente en dichas cartas.
Si los temas son complejos, el análisis del contenido es más difícil, especialmente si se complica al incluirse más de una oración simple.
- 3) *El ítem.* Tal vez es la unidad de análisis más utilizada y puede definirse como la unidad total empleada por los productores del material simbólico (Berelson 1952). Ejemplos de ítems pueden ser un libro, una editorial, un programa de radio o televisión, un discurso, una ley, un comercial, una carta amorosa, una conversación telefónica, una canción o la respuesta a una pregunta abierta. En este caso lo que se analiza es el material simbólico total.
- 4) *El personaje.* Un individuo, un personaje televisivo, un líder histórico, etcétera. Aquí lo que se analiza es el personaje.
- 5) *Medidas de espacio-tiempo.* Son unidades físicas como el centímetro-columna (por ejemplo, en la prensa), la línea (en escritos), el minuto (en una conversación o en radio), el periodo de 20 minutos (en una interacción), el cuadro (en televisión), cada vez que se haga una pausa (en un discurso).

Estas unidades se enclavan, colocan o caracterizan en categorías, esto podría representarse de la siguiente manera:

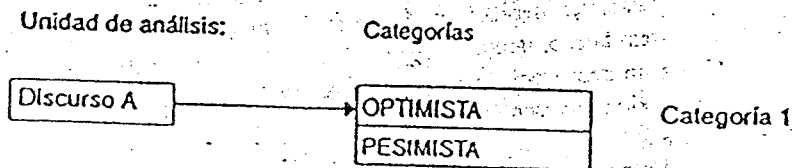


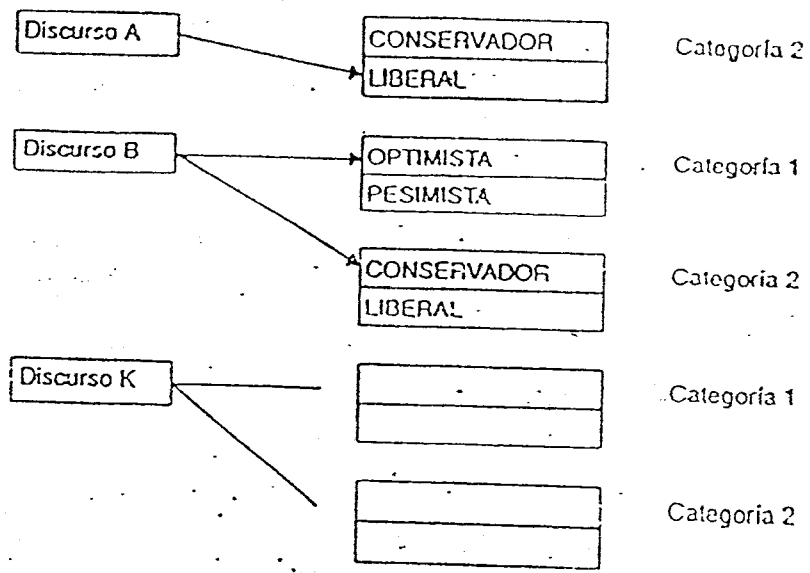
Respecto a la pregunta: ¿qué unidad debe seleccionarse? Esto depende de los objetivos y preguntas de investigación. Sin embargo, Berelson (1952) sugiere lo siguiente:

- a) En un solo estudio se pueden utilizar más de una unidad de análisis.
- b) Los cálculos de palabras y las unidades amplias, como el ítem y las medidas de espacio-tiempo, son más adecuadas en los análisis que dan énfasis a asuntos definidos.
- c) Las unidades amplias y las más definidas son válidas para la aceptación o rechazo en una categoría.
- d) Las unidades amplias generalmente requieren de menos tiempo para su codificación que las unidades pequeñas, referidas a las mismas categorías y materiales.
- e) Debido a que los temas u oraciones agregan otra dimensión al asunto, la mayoría de las veces son más difíciles de analizar que las palabras y las unidades amplias.
- f) El tema es adecuado para análisis de significados y las relaciones entre éstos.

CATEGORÍAS

Las categorías son los niveles donde serán caracterizadas las unidades de análisis. Tal y como menciona Holsti (1968), son las "casillas o cajones" en las cuales son clasificadas las unidades de análisis. Por ejemplo, un discurso podría clasificarse como optimista o pesimista, como liberal o conservador. Un personaje de una caricatura puede clasificarse como bueno, neutral o malo. En ambos casos, la unidad de análisis es categorizada. Veámoslo esquemáticamente:

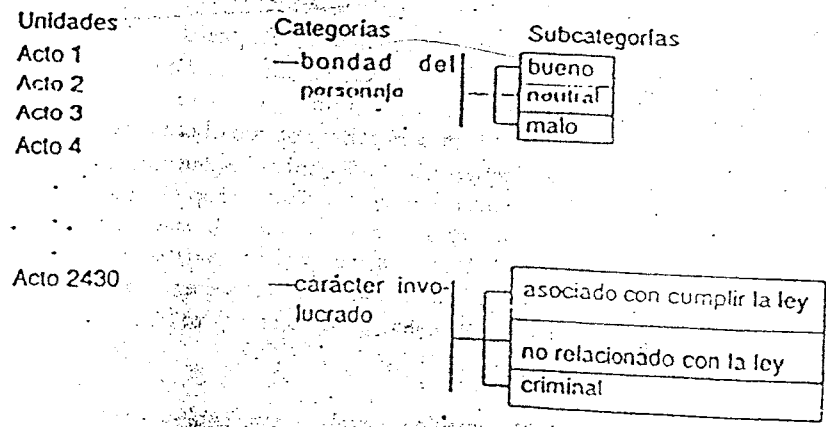




Es decir, cada unidad de análisis es categorizada o encasillada en uno o más sistemas de categorías. Por ejemplo, en un estudio citado por Krippendorff (1982) se analizaron 2 430 actos de violencia televisada, en cada acto el personaje principal (unidad de análisis) era categorizado como:

- Bueno, neutral o malo (sistema 1).
- Asociado con hacer cumplir la ley, no era relacionado con la ley o era presentado como un delincuente o criminal (sistema-2).

En este caso tenemos que la *unidad de análisis* es el comportamiento del personaje durante el acto televisivo, y las *categorías* eran dos: bondad del personaje y carácter involucrado. A su vez, las subcategorías de la bondad del personaje eran tres: bueno, neutral y malo. Y las subcategorías del carácter involucrado también eran tres: asociado con cumplir la ley, no relacionado con la ley y un criminal. Esto podría representarse así:



TIPOS DE CATEGORÍAS

Krippendorff (1982) señala cinco tipos de categorías:

- 1) *De asunto o tópico.* Las cuales se refieren a cuál es el asunto, tópico o tema tratado en el contenido (¿de qué trata el mensaje o la comunicación?).

EJEMPLO

- Análizar el último informe del Secretario o Ministro de Hacienda o Finanzas.

Categoría: Tono financiero
Subcategorías: Deuda
Impuestos
Planeación hacendaria
Inflación
Etc.

- 2) *De dirección.* Estas categorías se refieren a cómo es tratado el asunto (¿positiva o negativamente?, ¿favorable o desfavorable?, ¿nacionalista o no nacionalista?, etc.). Por ejemplo:

- Comparar la manera como dos noticiarios televisivos hablan de la posibilidad de una moratoria unilateral en el pago de la deuda externa de Latinoamérica.

Categoría: Tono en el tratamiento de la deuda externa.
Subcategorías: A favor de la moratoria unilateral.
En contra
Neutral

- 3) *De valores.* Se refieren a categorías que indican qué valores, intereses, metas, deseos o creencias son revelados. Por ejemplo:

- Al estudiar la compatibilidad ideológica de matrimonios, se podría analizar la ideología de cada cónyuge pidiéndoles un escrito sobre temas que puedan reflejar valores (sexo, actitud hacia la pareja, significado del matrimonio).

Categoría: Ideología del esposo.
Subcategorías: Muy tradicional
Más bien tradicional
Neutral
Más bien liberal
Muy liberal

- 4) *De receptores.* Estas categorías se relacionan con el destinatario de la comunicación (¿a quién van dirigidos los mensajes?).

EJEMPLO

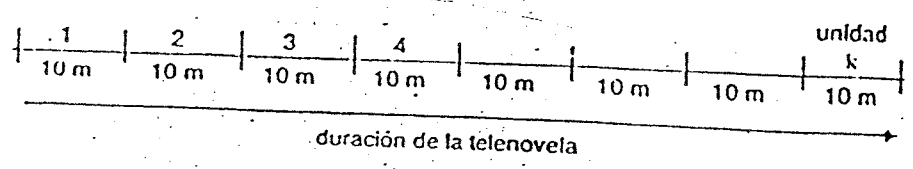
• Analizar a quiénes se dirige más un líder sindical en sus declaraciones a los medios de comunicación durante un periodo determinado.

- Categoría: Receptores a quienes dirige el mensaje.
- Subcategorías: Opinión pública en general
- Presidente
- Gabinete económico
- Gabinete agropecuario
- Gobierno en general
- Sector empresarial
- Obreros afiliados a su sindicato
- Obreros no afiliados a su sindicato
- Obreros en general (afiliados y no afiliados)
- Etc.

- 5) *Físicas.* Son categorías para ubicar la posición y duración o extensión de una unidad de análisis. De posición pueden ser por ejemplo la sección y página (en prensa), el horario (en televisión y radio). De duración, los minutos (en una interacción, un comercial televisivo, un programa de radio, un discurso), los centímetros/columna (en prensa), los cuadros en una película, etc. No se debe confundir las medidas de espacio-tiempo con las categorías físicas. Las primeras son unidades de análisis, las segundas constituyen categorías.

EJEMPLO

Cada periodo de 10 minutos de una telenovela lo voy a considerar la unidad de análisis.



- La unidad 1 la coloco en categorías.
- La unidad 2 la coloco en categorías.
- La unidad k la coloco en categorías.

Los minutos (cada 10) los considero una unidad.

Pero puedo tener categorías de tiempo. Por ejemplo, para analizar comerciales en radio:

- Categoría: Duración del comercial.
- Subcategorías: 10 segundos o menos
- 11-20 segundos
- 21-30 segundos
- Más de 30 segundos.

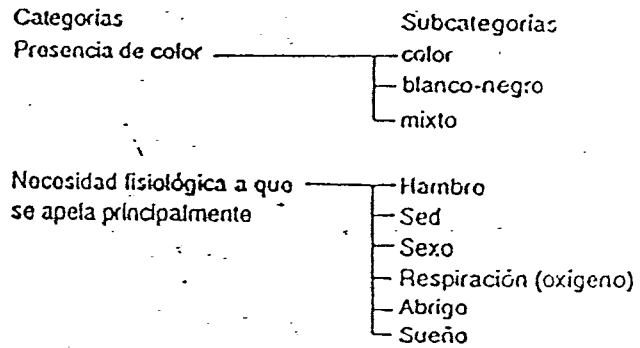
En ambos casos se usa el tiempo, pero en el primero como unidad y en el segundo como categoría. Es muy diferente.

REQUISITOS DE LAS CATEGORÍAS

En un análisis de contenido se suelen tener varias categorías, pero éstas deben cumplir con los siguientes requisitos:

- 1) Las categorías y subcategorías deben ser *exhaustivas*. Es decir, abarcar todas las posibles subcategorías de lo que se va a codificar. Por ejemplo, la categoría "Ideología del esposo" no podría prescindir de la subcategoría "neutral".
- 2) Las subcategorías deben ser *mutuamente excluyentes*, de tal manera que una unidad de análisis puede caer en una y sólo una de las subcategorías de cada categoría. Por ejemplo, un personaje no puede ser "bueno" y "malo" a la vez.

Aunque con respecto a las categorías no siempre son mutuamente excluyentes. Por ejemplo, al analizar comerciales televisivos podríamos tener —entre otras categorías— las siguientes:



Una unidad de análisis (un comercial) puede caer en una subcategoría de "Presencia de color" y en otra subcategoría de la categoría "Necesidad fisiológica a que se apela" (color y sed). Pero no puede caer en dos subcategorías de la misma categoría "hambre" y "sed", a menos que generáramos la subcategoría "hambre y sed":

Hambre y sed

Desde luego, en ciertos casos especiales, puede interesar al analista de contenido un sistema de categorías donde éstas no sean mutuamente excluyentes. Pero no es lo normal.

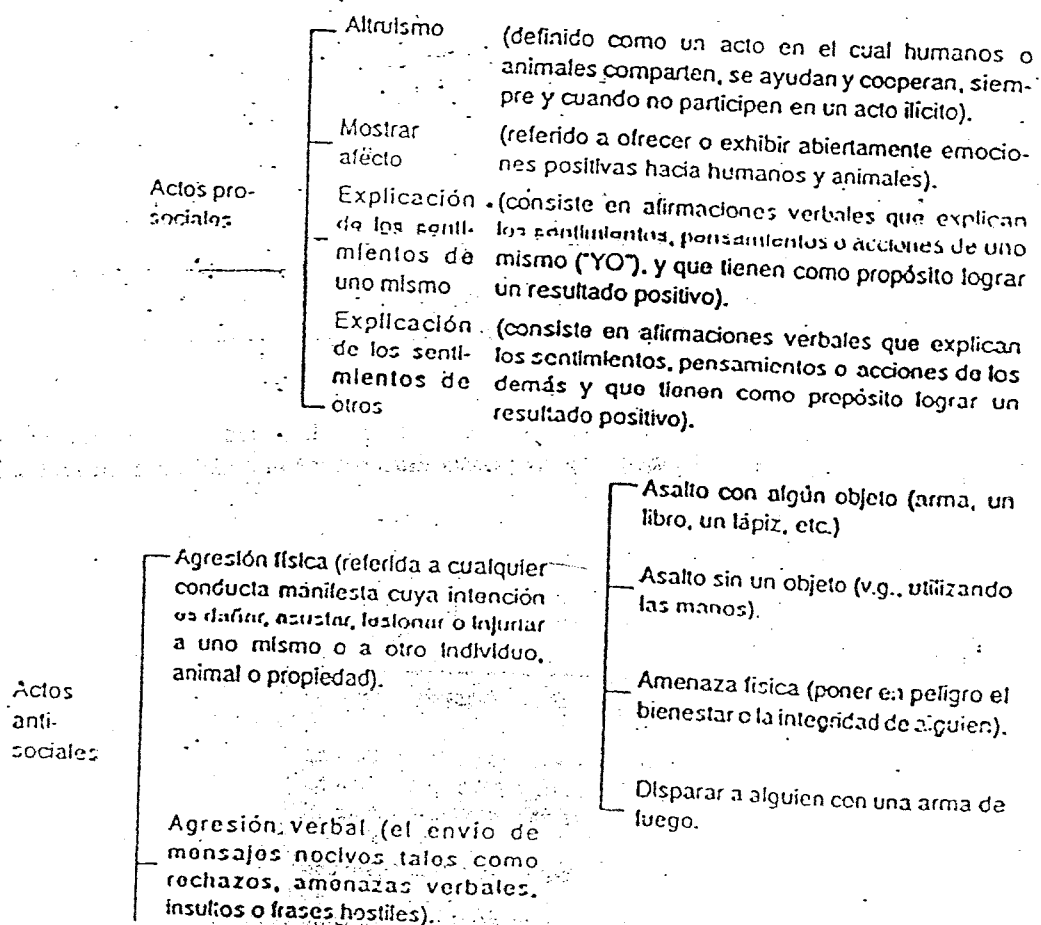
- 3) Las categorías y subcategorías deben derivarse del marco teórico y una profunda evaluación de la situación.

EJEMPLO DE UN ANÁLISIS DE CONTENIDO

Para ejemplificar el análisis de contenido y específicamente la generación de categorías acudiremos a un estudio de Greenberg, Edison, Korzeny, Fernández-Collado y Atkin (1980). El estudio consistió en un análisis de contenido de las series televisadas por las tres grandes cadenas de los Estados Unidos: ABC, CBS y NBC. Se analizaron diversos programas durante tres periodos de 1975 a 1978, para evaluar el grado en que la televisión norteamericana mostraba actos prosociales y antisociales como medida de la violencia televisada. Las categorías y subcategorías eran las siguientes:

FIGURA 9.18

EJEMPLO DE CATEGORÍAS



Robo (la acción de tomar intencional y deliberadamente una propiedad que pertenece a otra persona o institución, sin derecho o permiso).
 Engaño (engaño intencional de alguien con el propósito de provocar el detrimento de un individuo, grupo o institución (v.g., fraude, mentiras, asumir la identidad de otros).

El estudio consideró como unidad de análisis a la conducta, cada vez que una conducta se presentaba era codificada.

Cuando se crean las categorías, éstas deben ser definidas con precisión y es necesario explicitar qué se va a comprender en cada caso y qué habrá de excluirse.

El análisis de contenido consiste en asignar cada unidad a una o más categorías. De hecho, el producto de la codificación son frecuencias de cada categoría. Se cuenta cuántas veces se repite cada categoría o subcategoría (cuántas unidades de análisis entraron en cada una de las categorías). Por ejemplo, Greenberg et al. (1980, pp.113) encontraron los resultados que se muestran en la Tabla 9.5.

TABLA 9.5

EJEMPLO DE LOS RESULTADOS DE LA CODIFICACIÓN DE ACUERDO AL ESTUDIO DE GREENBERG et al. (1980)

	Año 1*		Año 2*		Año 3*	
	f	%	f	%	f	%
I. Agresión física						
A. Asalto con un objeto	466	(15.7)	248	(10.8)	370	(13.6)
B. Asalto sin un objeto	111	(3.7)	159	(6.9)	177	(6.5)
C. Amonaza física	180	(6.1)	233	(10.1)	135	(5.0)
D. Disparar	106	(3.6)	75	(3.2)	74	(2.7)
E. Otras	128	(4.3)	171	(7.4)	130	(4.8)
II. Agresión verbal	1 629	(55.0)	1 099	(47.6)	1 464	(54.0)
III. Robo	61	(2.1)	72	(3.1)	44	(1.6)
IV. Engaño	283	(9.5)	251	(10.9)	319	(11.8)
Total	2 964		2 308	(100.0)	2 713	(100.0)
Horas analizadas	(68.5)		(58)		(63)	

* Año 1 incluyó de octubre de 1975 hasta que se grabó un episodio de cada una de las series existentes (1976), año 2 igual pero en 1976-1977 y Año 3 igual pero de 1977-1978.

f = número de casos o unidades.

% = porcentajes.

¿CUÁLES SON LOS PASOS PARA LLEVAR A CABO EL ANÁLISIS DE CONTENIDO?

Ya hemos mencionado tres:

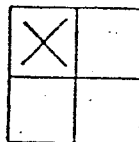
1. Definir con precisión el universo y extraer una muestra representativa.
2. Establecer y definir las unidades de análisis.
3. Establecer y definir las categorías y subcategorías que representen a las variables de la investigación.

Los demás pasos serían:

4. **Seleccionar a los codificadores.** Los codificadores son las personas que habrán de asignar las unidades de análisis a las categorías. Deben ser personas con un nivel educativo profesional (estudiantes a nivel de licenciatura como mínimo).
5. **Elaborar las hojas de codificación.** Estas hojas contienen las categorías y los codificadores anotan en ellas cada vez que una unidad entra en una categoría o subcategoría.

En la figura 9.19 se muestran los elementos de una hoja de codificación.

Las *hojas de codificación* pueden incluir elementos más precisos para describir el material. Por ejemplo, para analizar artículos de periódicos: Nombre del periódico, fecha de publicación, tipo de artículo (reportaje, entrevista, editorial, artículo de fondo, carta, inserción pagada, artículo general), signatario del artículo, sección donde se publicó, página, volumen de publicación (año, número o equivalente), ubicación,



tamaño (en centímetros/

columna), nombre del codificador, día de codificación, etc.

Una hoja puede servir para una o varias unidades, según que nos interese o no tener datos específicos de cada unidad (normalmente se prefiere tener la información por unidad).

Asimismo, las *categorías* y *subcategorías* deben estar *codificadas* con sus respectivos valores numéricos. En la hoja de codificación de la figura 9.19, 1 significa acto prosocial y 2 acto antisocial (y 1.1.0 es "altruismo", 1.2.0 "mostrar afecto", etcétera). Tres cifras en cada caso porque como se verá en el apartado sobre codificación, una variable debe tener el mismo número de columnas para todas sus categorías.

6. **Proporcionar entrenamiento de codificadores.** Este entrenamiento incluye que los codificadores se familiaricen y compenetren con las variables, comprendan las categorías y subcategorías y entiendan las definiciones de ambas. Además, debe capacitarse a los codificadores en la manera de codificar y debe discutirse ampliamente con ellos las diferentes condiciones en que puede manifestarse o estar presente cada categoría y subcategoría. Asimismo, los codificadores deben comprender completamente en qué consiste la unidad de análisis.

ELEMENTOS DE UNA HOJA DE CODIFICACIÓN

Indicador de quién codificó número, iniciales o letras

Fecha en que se codificó el material

Frecuencias (veces que se repite cada categoría)

Descripción del material que se analizará (discurso, nota periódica, sesión terapéutica, etc.)

CODIFICADOR: 1		MATERIAL A ANALIZAR:		CAPÍTULO DE LA SERIE "ALF" (CBS)	
FECHA: 29-XI-88		DURACIÓN: 30 MINUTOS			
CATEGORÍAS			FRECUENCIAS		TOTALES
ACTOS PROSOCIALES	ALTRUISMO (1.1.0.)			2	
	MOSTRAR AFECTO (1.2.0.)			5	
	EXPLICACIÓN DE LOS SENTIMIENTOS DE UNO (1.3.0.)			3	
	EXPLICACIÓN DE LOS SENTIMIENTOS DE OTROS (1.4.0.)			0	
ACTOS ANTISOCIALES	AGRESIÓN FÍSICA (2.1.)	Asalto con objeto (2.1.1.)			0
		Asalto sin objeto (2.1.2.)			0
		Amenaza física (2.1.3.)			0
		Disparar (2.1.4.)			0
		Otros (2.1.5.)			0
	AGRESIÓN VERBAL (2.2.0.)			3	
ROBO (2.3.0.)			0		
ENGANO (2.4.0.)			1		
TOTAL				14	

Suma de frecuencias en cada categoría o subcategoría

Al reverso pueden solicitarse comentarios del codificador.

Total de frecuencias o unidades

Al reverso pueden solicitarse comentarios del codificador.

Total de frecuencias o unidades

7. *Calcular la confiabilidad de los codificadores.* Una vez que se lleva a cabo el entrenamiento, los codificadores realizan una codificación provisional de una parte representativa del material (el mismo material para todos los codificadores), para ver si existe consenso entre ellos. Si no hay consenso no puede efectuarse un análisis de contenido confiable.

Para lo anterior se calcula la confiabilidad de cada codificador (individual) y la confiabilidad entre codificadores.

El cálculo de la *confiabilidad individual* de los codificadores depende de si tenemos uno o varios de éstos.

A. *Si se dispone de un solo codificador* (porque el material es reducido), se observan las diferencias de la codificación del mismo mensaje hecha por el codificador en dos tiempos diferentes. Si las diferencias son muy pequeñas, el codificador es individualmente confiable. Este tipo de confiabilidad es llamada "confiabilidad intra-codificador". La cual mide la estabilidad de la prueba y re-prueba de un codificador a través del tiempo.

Otro método consistiría en que el codificador trabaje una parte representativa del material y después aplicar a su codificación (resultados) la siguiente fórmula:

$$\text{Confiabilidad individual} = \frac{\text{Número de unidades de análisis catalogadas correctamente por el codificador}}{\text{Número total de unidades de análisis}}$$

Suponiendo que un mensaje conste de 20 unidades y se legren correctamente 20, la confiabilidad será de $1 \left(\frac{20}{20} \right)$, que es el máximo de confiabilidad. Si el codificador sólo pudo codificar adecuadamente 15 de los 20, la confiabilidad sería de $0.75 \left(\frac{15}{20} \right)$.

B. *Si se dispone de varios codificadores*, la confiabilidad individual puede determinarse así: Se les pide a todos los codificadores que codifiquen el mismo material, se toman los resultados de todos menos los de uno y se compara la codificación de éste contra la del resto. Así se procede con cada codificador.

También puede aplicarse a todos los codificadores la fórmula mencionada para calcular la confiabilidad individual y quien se distancie del resto se considera un caso poco confiable.

EJEMPLO

Codificador A	Codificador B	Codificador C	Codificador D
0.89	0.93	0.92	0.67

El codificador "D" tiene baja confiabilidad.

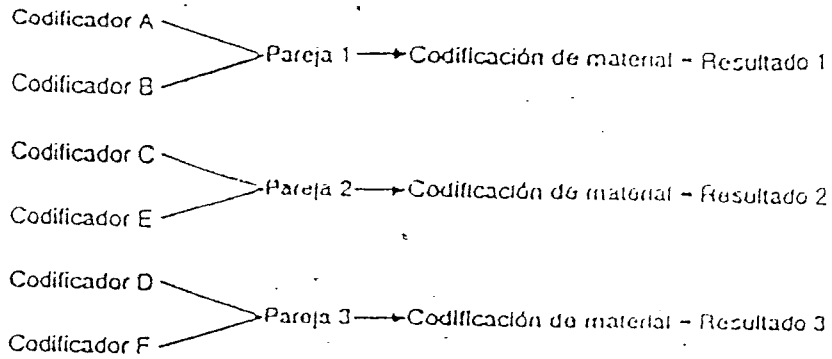
El cálculo de la *confiabilidad intercodificadores* se realiza por pares de codificadores (parejas). Se pide a cada pareja formada que codifique el material, se comparan los resultados obtenidos por las parejas, se cuenta el número de acuerdos entre las parejas, se determina el número de unidades de análisis y se aplica la siguiente fórmula:

$$\text{Confiabilidad entre parejas} = \frac{\text{Número total de acuerdos entre dos parejas}}{\text{Número total de unidades de análisis codificadas}}$$

Después se suman los resultados de esta fórmula y se divide entre el número de comparaciones, que depende del número de parejas.

Vémoslo con un ejemplo:

Las parejas codifican el material:



Se determina el número de acuerdos entre las parejas (un acuerdo consiste en que dos parejas codifican en la misma categoría a una misma unidad de análisis):

Pareja	Número de unidades de análisis codificadas	Número de acuerdos entre parejas
1	18	Entre parejas 1 y 2 - 17
2	18	Entre parejas 1 y 3 - 16
3	17	Entre parejas 2 y 3 - 16

Se aplica la fórmula de confiabilidad entre parejas

$$C_{1y2} = \frac{17}{18} = 0.94$$

$$C_{1y3} = \frac{16}{18} = 0.88$$

$$C_{2y3} = \frac{16}{18} = 0.88$$

Debe observarse que no hubo consenso total entre cuántas unidades de análisis podían distinguirse en el material (la pareja 3 distinguió 17 unidades y las parejas 1 y 2 distinguieron 18). En este caso se toma en cuenta para la fórmula de confiabilidad entre parejas, el máximo número de unidades de análisis codificadas por alguna de las parejas. Si fuera:

Pareja	Número de unidades de análisis codificadas	Número de acuerdos entre parejas
A	25	21
B	22	

so toma el máximo

La fórmula sería $C_{AB} = \frac{21}{25} = 0.84$

Se obtiene la *confiabilidad total* (que es la suma de las confiabilidades de las parejas sobre el número de comparaciones). En nuestro ejemplo:

$$C_T = \frac{C_{1y2} + C_{1y3} + C_{2y3}}{3}$$

$$C_T = \frac{0.94 + 0.88 + 0.88}{3} = 0.90$$

No es conveniente tolerar una *confiabilidad* menor que 0.85 (ni total ni en las parejas) y de ser posible debe superar el 0.89. Al igual que con otros instrumentos de medición, la confiabilidad oscila entre 0 (nula confiabilidad) y 1 (confiabilidad perfecta).

En el análisis de contenido una *baja confiabilidad puede deberse* o a las categorías y/o unidades de análisis no han sido definidas con claridad y precisión, o a un deficiente entrenamiento o a inhabilidad de los codificadores. Cuando se detecta una baja confiabilidad debe detectarse y corregirse el problema.

Asimismo, es conveniente calcular la confiabilidad a la mitad de la escala (con el material codificado) y al finalizar ésta.

8. *Efectuar la codificación.* Lo que implica contar las frecuencias de repeticiones en las categorías (número de unidades que entran en cada categoría)
9. *Vaciar los datos de las hojas de codificación y obtener totales por categoría.*
10. *Realizar los análisis estadísticos apropiados.*

9.6.4. Observación

¿QUÉ ES Y PARA QUÉ SIRVE LA OBSERVACIÓN?

La observación consiste en el registro sistemático, válido y confiable de comportamientos o conducta manifiesta. Puede utilizarse como instrumento de medición en muy diversas circunstancias. Haynes (1978) menciona que es un método más utilizado por quienes están orientados conductualmente. Puede servir para determinar la aceptación de un grupo respecto a su profesor, analizar conflictos familiares, conductas masivas (v.g., la violencia en los estadios de fútbol), la aceptación de un producto en un supermercado, el comportamiento de deficientes mentales, etcétera.

Como método para recolectar datos es muy similar al análisis de contenido. De hecho, éste es una forma de observación del contenido de comunicaciones. En este apartado algunos conceptos sólo serán mencionados, pues han sido tratados en el apartado sobre análisis del contenido.

PASOS PARA CONSTRUIR UN SISTEMA DE OBSERVACIÓN

Los pasos para construir un sistema de observación son:

1. *Definir con precisión el universo de aspectos, eventos o conductas a observar.* Por ejemplo, si nuestro interés es observar los recursos con que cuenta el

escuelas de un distrito escolar debemos definir lo que concebimos como "recurso escolar". Un universo podría ser el comportamiento verbal y no verbal de un grupo de alumnos durante un semestre. Otro universo sería las conductas de un grupo de trabajadores durante sus sesiones en círculos de calidad o equipos para la calidad, en un periodo de un año. O bien las conductas agresivas de un grupo de esquizofrénicos en sesiones terapéuticas.

2. *Extraer una muestra representativa de los aspectos, eventos o conductas a observar.* Un repertorio suficiente de conductas para observar.
3. *Establecer y definir las unidades de observación.* Por ejemplo, cada vez que se presenta una conducta agresiva, cada minuto se analizará si el alumno está o no atento a la clase, durante dos horas al día (7:00 a 9:00 horas), el número de personas que leyeron el tablero de avisos de la compañía, etcétera. El concepto de unidad de análisis es el mismo que en el análisis de contenido solamente que en la observación se trata de conductas, eventos o aspectos.
4. *Establecer y definir las categorías y subcategorías de observación.* Estas categorías son similares a las definidas para el análisis de contenido. Y la observación también consiste en asignar unidades a categorías y subcategorías de observación.

EJEMPLO DEL CUARTO PASO

En el caso del estudio citado —al hablar de la manipulación de variables independientes en experimentos en la página 114— para probar la hipótesis: "A mayor grado de información sobre la deficiencia mental que el sujeto normal maneje, mostrará menor evitación en la interacción con el deficiente mental" (Naves y Poplawsky, 1984): Las unidades de análisis eran cada 10 segundos. La interacción entre la persona normal y el actor que hacía el papel de "deficiente mental" duraba tres minutos. La variable dependiente fue "evitación de la interacción" y las categorías fueron cuatro (Naves y Poplawsky, 1984, pp. 107-109):

1. **DISTANCIA FÍSICA:** se refiere a si el sujeto experimental aumenta o disminuye su distancia hacia el interlocutor a partir de la distancia que inicialmente debía ocupar; esta distancia inicial estuvo delimitada por los asientos que el actor y el sujeto debían ocupar y, según la teoría, es la distancia en la que dos extraños en una situación de comunicación, pueden interactuar cómodamente. Las dimensiones que esta variable adquiere son el acercarse (afiliación) con valor de "1" o el alejarse (evitación) del actor (deficiente mental) con valor de "0", mediante inclinaciones corporales o bien modificando por completo su distancia.
2. **MOVIMIENTOS CORPORALES QUE DENOTAN TENSION:** esta variable se orienta a captar los movimientos que el sujeto está realizando como índice de tensión (evitación) con valor de "0" o de relajación (afiliación) que experimenta, con valor de "1". En esta variable específicamente se analizan movimientos de pies y piernas a un ritmo acelerado, ademanes con brazos y

manos (como el estarse rascando, picando, etc.) y la postura en general del sujeto.

3. CONDUCTA VISUAL DEL SUJETO: que según lo estipulado en esta investigación adquiere dos dimensiones:
 - a) dirigida hacia el interlocutor (afiliación), con valor de "1".
 - b) dirigida hacia cualquier otra parte (evitación), con valor de "0".
4. CONDUCTA VERBAL: este indicador está compuesto por el contenido verbal del discurso del sujeto hacia el deficiente y se orienta primordialmente al formato del discurso; incluye dos modalidades:
 - a) frases u oraciones completas (afiliación), con valor de "1".
 - b) frases dicótomas y silencios (evitación), con valor de "0".

La modalidad de frases dicótomas incluye respuestas monosilábicas (sí, no), murmulos, sonidos guturales, etc., y los silencios que se interpretan como respuestas dicótomas (respuesta de evitación) son los silencios no naturales en el discurso, aquellos en los que expresamente el sujeto se queda un periodo en silencio. La conducta verbal se mide a través del diálogo que sostenga el sujeto para con el deficiente mental; es decir, en respuesta al guión que el actor interpreta para con cada sujeto (que es idéntico para todos) y en las intervenciones que el propio sujeto realice. En un principio se pensó que además de medir la conducta verbal en cuanto a su formato, convendría medirla en cuanto a su contenido también; es decir, si las frases emitidas por él en respuesta a lo expresado por el deficiente revelaban un contenido positivo o negativo. Esta modalidad no fue incluida por la dificultad que presenta el obtener una medición objetiva.

Finalmente, cabe establecer que en los cuatro indicadores (variables) adoptados para medir la evitación en la interacción se establecieron dimensiones cuya medición fuese dicotómica; es decir, que las opciones de respuesta para cada variable únicamente pudiesen ser codificadas bajo la escala de 0 — 1; esto es, evitación o no evitación. La razón por la cual dichas variables no adquieren más opciones de respuesta obedece a la dificultad por detectar conductas de evitación tanto verbales como no verbales.

LAS SUBCATEGORÍAS PUEDEN SER ESCALAS DE ACTITUDES

Al establecer las subcategorías, éstas pueden ser escalas del tipo Likert, Guttman o diferencial semántico (recuérdese la sección 2.6.1).

EJEMPLO

Categoría
—Atención del alumno

Subcategorías

—	Elevada	(3)
—	Media	(2)
—	Baja	(1)
—	Nula	(0)

FIGURA 9.20

EJEMPLO DE UNA HOJA DE CODIFICACIÓN PARA OBSERVAR CONDUCTA

Nombre _____

Grupo _____

Fecha _____

CATEGORÍA	SUBCATEGORÍA	0'	10'	20'	30'	40'	50'	1'	10'	20'	30'	40'	50'	2'	10'	20'	30'	40'	50'	J	
DISTANCIA FÍSICA	ALEJAMIENTO (0)																				
	ACERCAMIENTO (2)																				
	ESTÁTICO (1)																				
MOVIMIENTOS CORPORALES																					
	TENSIÓN (0)																				
	RELAJACIÓN (2)																				
	NINGUNO (1)																				
CONDUCTA VISUAL																					
	AL SUJETO (1)																				
	A OTRA PARTE (0)																				
CONDUCTA VERBAL																					
	F. COMPLETAS (1)																				
	F. DICOTOMAS (0)																				
	SILENCIOS (0)																				

Nota: Se agregaron las subcategorías "estático" en la categoría "distancia física" y "ninguno" en la categoría "movimientos corporales".

8. *Proporcionar entrenamiento de codificadores* (en las variables, categorías, subcategorías, unidades de análisis y el procedimiento de codificar, así como sobre las diferentes maneras como puede manifestarse una categoría o subcategoría de conducta).
9. *Calcular la confiabilidad de los observadores* (intra-observador e interobservadores). Los procedimientos y fórmulas pueden ser las mismas que las vistas en el apartado sobre el análisis de contenido, lo único que cambia es la palabra "codificador(es)", "codificación", "codificada(s)"; por "observador(es)", "observación", "observada(s)".

Por ejemplo:

$$\text{Confiabilidad individual} = \frac{\text{Número de unidades de análisis catalogadas correctamente por el observador}}{\text{Número total de unidades de análisis}}$$

$$\text{Confiabilidad entre parejas} = \frac{\text{Número total de acuerdos entre dos parejas}}{\text{Número total de unidades de análisis observadas}}$$

Haynes (1978, p. 160) proporciona otra fórmula para calcular la confiabilidad entre observadores o el grado de acuerdo interobservadores (λ_o).

$$\lambda_o = \frac{I_a}{I_a + I_d}$$

Donde "Ia" es el número total de acuerdos entre observadores e "Id" es el número total de desacuerdos entre observadores. Un "Acuerdo" es definido como la codificación de una unidad de análisis en una misma categoría por distintos observadores. Se interpreta como cualquier coeficiente de confiabilidad (0 a 1).

10. Llevar a cabo la codificación por observación.
11. Vaciar los datos de las hojas de codificación y obtener totales para cada categoría.
12. Realizar los análisis apropiados.

TIPOS DE OBSERVACIÓN

La observación puede ser *participante* o *no participante*. En la primera, el observador interactúa con los sujetos observados y en la segunda no ocurre esta interacción. Por ejemplo, un estudio sobre las conductas de aprendizaje de niños autistas, en donde una instructora interactúa con los niños y al mismo tiempo codifica.

VENTAJAS DE LA OBSERVACIÓN

Tanto la observación como el análisis de contenido tienen varias ventajas:

- 1) Son técnicas de medición no obstrusivas. En el sentido que el instrumento de medición no "estimula" el comportamiento de los sujetos (las escalas de actitud y los cuestionarios pretenden "estimular" una respuesta a cada ítem). Los métodos no obstrusivos simplemente registran algo que fue estimulado por otros factores ajenos al instrumento de medición.
- 2) Aceptan material no estructurado.
- 3) Pueden trabajar con grandes volúmenes de datos (material).

9.6.5. Pruebas e inventarios estandarizados

¿QUÉ SON LAS PRUEBAS ESTANDARIZADAS?

En la actualidad existe una amplia diversidad de pruebas e inventarios desarrollados por diversos investigadores para medir un gran número de variables. Estas pruebas tienen su propio procedimiento de aplicación, codificación e interpretación, y se encuentran disponibles en diversas fuentes secundarias y terciarias, así como en centros de investigación y difusión del conocimiento. Hay pruebas para medir habilidades y aptitudes (v.g., habilidad verbal, razonamiento, memoria, inteligencia, percepción, habilidad numérica), la personalidad, los intereses, los valores, el desempeño, la motivación, el aprendizaje, el clima laboral en una organización, etcétera. También se puede disponer de pruebas clínicas para detectar conducta anormal, pruebas para seleccionar personal, pruebas para conocer las percepciones y/o opiniones de las personas respecto a diversos tópicos, pruebas para medir la autoestima y —en fin— otras muchas variables del comportamiento.⁴⁰

El problema en el uso de estas pruebas es que la mayoría ha sido desarrollada en contextos muy diferentes al latinoamericano, y en ocasiones su utilización puede ser inadecuada, inválida y poco confiable. Cuando se utilice como instrumento de medición una prueba estandarizada es conveniente que se seleccione una prueba desarrollada o adaptada por algún investigador para el mismo contexto de nuestro estudio y que sea válida y confiable (debemos tener información a este respecto). En el caso de que elijamos una prueba diseñada en otro contexto, es necesario adaptarla y aplicar pruebas piloto para calcular su validez y confiabilidad, así como ajustarla a las condiciones de nuestra investigación. El instrumento o prueba debe demostrar que es válido y confiable para el contexto en el cual se va a aplicar.

Un tipo de pruebas estandarizadas bastante difundido lo constituyen las "pruebas proyectivas", las cuales presentan estímulos a los sujetos para que respondan a ellos; después se pueden analizar las respuestas tanto cuantitativamente como cualitativamente y se interpretan. Estas pruebas miden proyecciones de los sujetos, como por ejemplo, la personalidad.

Dos pruebas proyectivas muy conocidas son el *Test de Rorschach* (que presenta a los sujetos manchas de tinta en tarjetas o láminas blancas numeradas y éstos relatan sus asociaciones e interpretaciones en relación a las manchas) y el *Test de Apercepción Temática* (que con un esquema similar al de Rorschach presenta a los sujetos cuadros que evocan narraciones o cuentos y las personas deben elaborar una interpretación).

Hastantes pruebas estandarizadas (v.g., las proyectivas) requieren de un entrenamiento considerable y un conocimiento profundo de las variables por parte del investigador que habrá de aplicarlas e interpretarlas. No pueden aplicarse con superficialidad e indiscriminadamente. La manera de aplicar, codificar, calificar e interpretar las pruebas estandarizadas es tan variada como los tipos existentes.

⁴⁰ Para conocer la diversidad de estas pruebas y sus aplicaciones se recomienda Anastasi (1972), Thorndike y Hagen (1980), Cronbach (1984) y Nunnally (1970). Son obras clásicas sobre mediciones y el manejo de pruebas estandarizadas.

9.6.6. Sesiones en profundidad

¿QUÉ SON LAS SESIONES EN PROFUNDIDAD?

Un método de recolección de datos cuya popularidad ha crecido son las sesiones en profundidad. Se reúne a un grupo de personas y se trabaja con éste en relación a las variables de la investigación. Pueden realizarse una o varias reuniones. El procedimiento usual es el siguiente.

PASOS PARA REALIZAR LAS SESIONES DE GRUPO

1. Se define el tipo de personas que habrán de participar en la sesión o sesiones.
2. Se detectan personas del tipo elegido.
3. Se invita a estas personas a la sesión o sesiones.
4. Se organizan la sesión o sesiones. Cada sesión debe efectuarse en un lugar confortable, silencioso y aislado. Los sujetos deben sentirse cómodos y relajados. Asimismo, es indispensable planear cuidadosamente lo que se va a tratar en la sesión o sesiones (desarrollar una agenda) y asegurar los detalles (aún las cuestiones más sencillas como el servir café y refrescos).
5. Se lleva a cabo cada sesión. El conductor debe ser una persona entrenada en el manejo o conducción de grupos y debe crear "rapport" en el grupo (clima de confianza). Asimismo, debe ser un individuo que no sea percibido como "distante" por los participantes de la sesión y tiene que propiciar la participación de todos. La paciencia es una característica que deberá tener.

Durante la sesión se pueden pedir opiniones, hacer preguntas, administrar cuestionarios, discutir casos, intercambiar puntos de vista, valorar diversos aspectos. Es conveniente que cada sesión se grabe en cinta o videocinta y después realizar análisis de contenido y observación. El conductor debe tener muy en claro la información o datos que habrán de recolectarse y evitar desviaciones del objetivo planteado.

6. Elaborar el reporte de sesión. El cual incluye principalmente datos sobre los participantes (edad, sexo, nivel educativo y todo aquello que sea relevante para el estudio), fecha y duración de la sesión, información completa del desarrollo de la sesión, actitud y comportamiento de los participantes hacia el conductor y la sesión en sí, resultados de la sesión y observaciones del conductor, así como una bitácora de la sesión.
7. Llevar a cabo la codificación y análisis correspondientes.

EJEMPLOS

Algunos ejemplos de la aplicación de este método podrían ser las sesiones en donde se evalúe a un nuevo producto, digamos un dulce. En estas sesiones se podría pedir a los participantes opiniones sobre el sabor, color, presentación, precio, cualidades, etcétera, del producto; discutir a fondo las propiedades,

evaluaciones y carencias del producto; administrarles una escala de actitudes o un cuestionario, y hacer preguntas abiertas a cada participante. O bien sesiones para analizar la popularidad de varios candidatos políticos, evaluar el servicio y la atención recibida en un supermercado, indagar la percepción de un grupo de estudiantes sobre la calidad de la enseñanza recibida o conocer la opinión de los sectores de una comunidad sobre una reforma electoral.

También se pueden organizar sesiones con diferentes tipos de la población y mixtas. Por ejemplo, los autores participaron en un estudio para evaluar un programa televisivo que acababa de "salir al aire" con un nombre-formato nuevos. Además de realizarse una encuesta telefónica se organizaron varias sesiones (algunas con amas de casa, otras con estudiantes, también con trabajadores, empleados de oficina, ejecutivos, profesores, publicistas y otros grupos tipo; así como reuniones donde participaban —por ejemplo— un ama de casa de más de 60 años, un ama de casa de 50 años, un ama de casa más joven, un empleado de una oficina pública, una secretaria, un profesor, un dependiente de supermercado, un ejecutivo y dos estudiantes.

En las sesiones se profundizó en el formato del programa (música, manejo de cámaras, duración, manejo de comerciales, sonido, manera de presentar invitados, etc.), los conductores, el contenido, etcétera.

Normalmente en las sesiones participan de ocho a quince personas. No debe excederse de un número manejable de sujetos. El formato y naturaleza de la sesión o sesiones depende del objetivo y las características de los participantes.

9.6.7. Otras formas de recolección de los datos

¿QUÉ OTRAS MANERAS EXISTEN PARA RECOLECTAR LOS DATOS?

En ocasiones puede acudirse a *archivos* que contengan los datos. Por ejemplo, podemos acudir a la alcaldía de algunas ciudades para solicitar datos relacionados con la violencia (si nuestra hipótesis fuera: "La violencia manifiesta en la ciudad de México es mayor que en la ciudad de Caracas"): número de asaltos, violaciones, robos a casa-habitación, asesinatos, etc. (datos por habitante, distrito y generales). También podríamos acudir a los hospitales y las diferentes procuradurías, etc. Otro ejemplo, sería consultar los archivos de una universidad y tomar los datos de inteligencia, personalidad u otras variables que nos interesen. Desde luego, a veces esta información no es accesible. En México hay un organismo que proporciona datos (incluso grabados en disco para computadora) sobre estadísticas nacionales, el Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Asimismo, pueden utilizarse datos recolectados por otros investigadores, a lo que se conoce como "*análisis secundario*". En este caso es necesario tener la certeza de que los datos son válidos y confiables, así como conocer la manera como fueron codificados. El intercambio de éstos es una práctica común entre investigadores. Además, existen métodos propios de las diferentes ciencias sociales como el *análisis de redes* para evaluar cómo se manifiesta la comunicación en un sistema social (quién se comunica con quién, quiénes distorsionan la información, cómo fluye la comuni-

- 1) Codificar las categorías de ítems, preguntas y categorías de contenido u observación no precodificadas.
- 2) Elaborar el libro de códigos.
- 3) Efectuar físicamente la codificación.
- 4) Grabar y guardar los datos en un archivo permanente.

Veamos cada paso con algunos ejemplos.

Codificar

Si todas las categorías fueron precodificadas y no se tienen preguntas abiertas, el primer paso no es necesario. Éste ya se efectuó.

Si las categorías no fueron precodificadas y se tienen preguntas abiertas, deben asignarse los códigos o la codificación a todas las categorías de los ítems, preguntas o de contenido u observación. Por ejemplo:

Pregunta no precodificada

¿Practica usted algún deporte por lo menos una vez a la semana?

Sí

No

Se codifica

1 - Sí

0 - No

Frase no precodificada

"Creo que estoy recibiendo un salario justo por mi trabajo"

() Totalmente de acuerdo

() De acuerdo

() Ni de acuerdo, ni en desacuerdo

() En desacuerdo

() Totalmente en desacuerdo

Se codifica

5 - Totalmente de acuerdo

4 - De acuerdo

3 - Ni de acuerdo, ni en desacuerdo

2 - En desacuerdo

1 - Totalmente en desacuerdo

Tratándose de preguntas abiertas ya se expuso cómo se codifican.

Libro de códigos

Una vez que están codificadas todas las categorías del instrumento de medición, se procede a elaborar el "libro de códigos".

El libro de códigos es un documento que describe la localización de las variables y los códigos asignados a los atributos que las componen (categorías y/o subcategorías) (Babbie, 1979). Este libro cumple con dos funciones: i) es la guía para el proceso de codificación y ii) es la guía para localizar variables e interpretar los datos durante el análisis (Babbie, 1979). El libro de códigos puede conducirnos a los significados de los valores de las categorías de las variables.

Los elementos de un libro de códigos son: *variable, pregunta / ítem / tema, categorías-subcategorías, columna(s)*.

Supongamos que tenemos una escala Likert con tres ítems (frases):

"La Dirección General de Impuestos Nacionales informa oportunamente sobre cómo, dónde y cuándo pagar los impuestos"

- (5) Muy de acuerdo
- (4) De acuerdo
- (3) Ni de acuerdo, ni en desacuerdo
- (2) En desacuerdo
- (1) Muy en desacuerdo

"Los servicios que presta la Dirección General de Impuestos Nacionales son en general muy buenos"

- (5) Muy de acuerdo
- (4) De acuerdo
- (3) Ni de acuerdo, ni en desacuerdo
- (2) En desacuerdo
- (1) Muy en desacuerdo

"La Dirección General de Impuestos Nacionales se caracteriza por la deshonestidad de sus funcionarios"

- (1) Muy de acuerdo
- (2) De acuerdo
- (3) Ni de acuerdo, ni en desacuerdo
- (4) En desacuerdo
- (5) Muy en desacuerdo

El libro de códigos sería el que se muestra en la figura 9.21.

FIGURA 9.21

EJEMPLO DE UN LIBRO DE CÓDIGOS CON UNA ESCALA DE ACTITUD TIPO LIKERT (TRES ÍTEMS)

VARIABLE	ÍTEM	CATEGORÍAS	CÓDIGOS	COLUMNA
— Actitud hacia la Dirección General de Impuestos Nacionales.	Frase 1 (informa)	— Muy de acuerdo	5	1
		— De acuerdo	4	
		— Ni de acuerdo, ni en desacuerdo	3	
		— En desacuerdo	2	
		— Muy en desacuerdo	1	
	Frase 2 (servicios)	— Muy de acuerdo	5	2
		— De acuerdo	4	
		— Ni de acuerdo, ni en desacuerdo	3	
		— En desacuerdo	2	
		— Muy en desacuerdo	1	
	Frase 3 (deshonestidad)	— Muy de acuerdo	1	3
		— De acuerdo	2	
		— Ni de acuerdo, ni en desacuerdo	3	
		— En desacuerdo	4	
		— Muy en desacuerdo	5	

En el caso del estudio por observación de Naves y Poplawsky (1984) (figura 9.20), el libro de códigos sería el que se muestra en la figura 9.22.

FIGURA 9.22

EJEMPLO DE UN LIBRO DE CÓDIGOS EN EL CASO DEL ESTUDIO DE NAVES Y POPLAWSKY (1984)⁴⁷

LIBRO DE CÓDIGOS				
VARIABLE	CATEGORÍAS	SUBCATEGORÍAS	CÓDIGOS	COLUMNAS
— Número de sujeto	— Sujetos	—	{01 a 30 {(hubo 30 {sujetos)	1 Y 2

⁴⁷ Desde luego, Naves y Poplawsky (1984) para las categorías de "conducta de evasión" obtenían esta codificación cada 10 segundos (cada unidad de análisis), y sumaban el número de 1 (unos) y lo transformaban en porcentajes. Aquí suponemos que toda la interacción con el deficiente mental puede categorizarse y subcategorizarse.

Tratamiento experimental	Grupo cultural		1	3
	Grupo socio-psicológico		2	
Conducta del excitación	Desafectación	Alejamiento	0	4
		Aproximamiento		
	Movimientos corporales	Estáticos		1
		Tensión		0
	Conducta visual	Relajación		2
		Ninguno		1
		Al sujeto		1
	Conducta verbal	A otra parte		0
		Frasas completas		1
	Codificador	Frasas dicotomas o silencios		0
LRE			1	
LEMM			2	
	CFT		3	
				6
				7
				8

Es decir, el libro de códigos es un manual para el investigador y los codificadores. Los cuestionarios contestados, las escalas aplicadas, las hojas de codificación, las pruebas respondidas o cualquier otro instrumento de medición administrado son transferidos a una matriz, la cual es el conjunto de datos simbólicos o numéricos producto de la aplicación del instrumento. Esta matriz es lo que habrá de analizarse. El apartado "columna" dentro del libro de códigos tiene sentido en la matriz; veamos por qué. La matriz tiene renglones y columnas; los renglones representan casos o sujetos en la investigación, las columnas son los lugares donde se registran los valores en las categorías o subcategorías. Esto podría esquematizarse así:

MATRIZ DE DATOS

	Columna 1	Columna 2	Columna 3	
Caso 1				Valores de categorías/subcategorías
Caso 2				
Caso 3				
Caso 4				
Caso k				

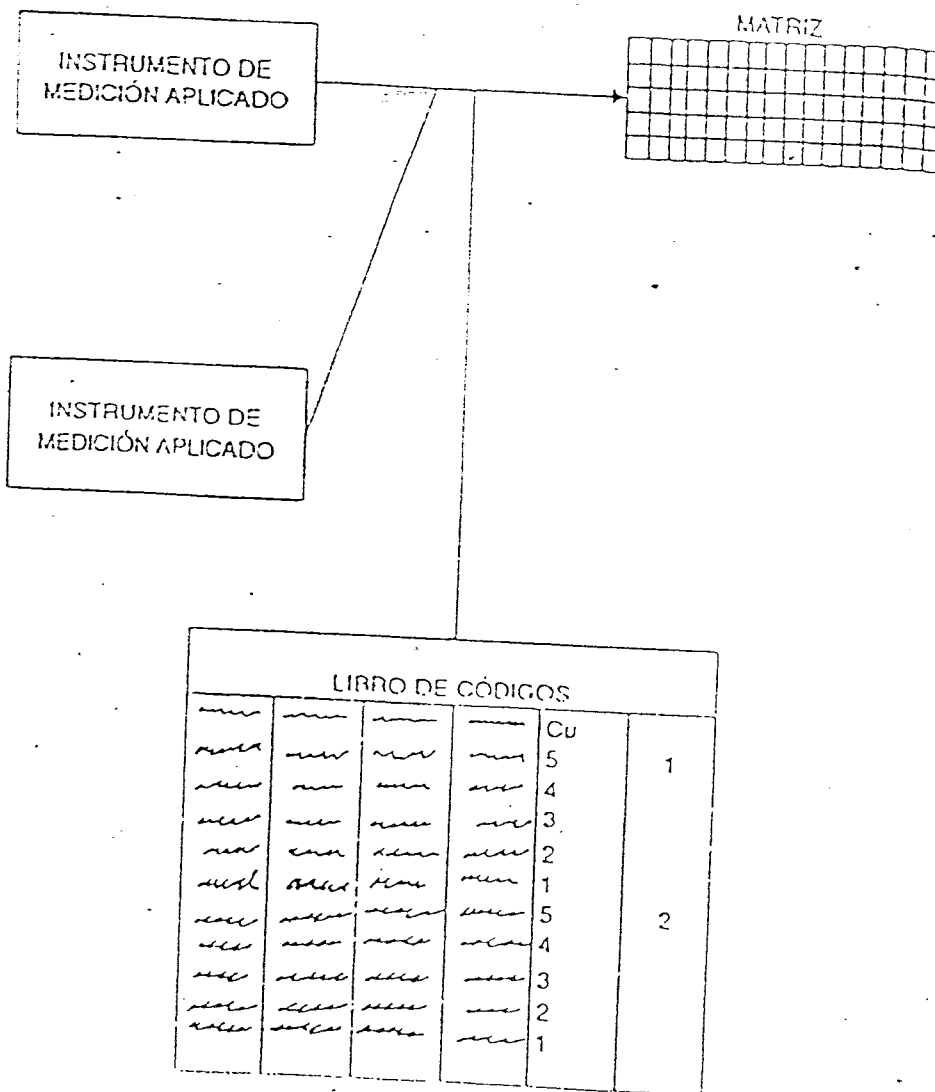
— Tratamiento experimental	— Grupo cultural		1	3
	— Grupo socio-psicológico		2	
— Conducta de evitación	— Distancia física	— Alejamiento	0	4
		— Acercamiento	2	
		— Estático	1	
	— Movimientos corporales	— Tensión	0	5
		— Relajación	2	
		— Ninguno	1	
	— Conducta visual	— Al sujeto	1	6
		— A otra parte	0	
	— Conducta verbal	— Frases completas	1	7
		— Frases dicótomas o silencios	0	
— Codificador	— LRE	—	1	8
	— LEMM	—	2	
	— CFT	—	3	

Es decir, el libro de códigos es un manual para el investigador y los codificadores. Los cuestionarios contestados, las escalas aplicadas, las hojas de codificación, las pruebas respondidas o cualquier otro instrumento de medición administrado son transferidos a una matriz, la cual es el conjunto de datos simbólicos o numéricos producto de la aplicación del instrumento. Esta matriz es lo que habrá de analizarse. El apartado "columna" dentro del libro de códigos tiene sentido en la matriz; veamos por qué. La matriz tiene renglones y columnas; los renglones representan casos o sujetos en la investigación, las columnas son los lugares donde se registran los valores en las categorías o subcategorías. Esto podría esquematizarse así:

MATRIZ DE DATOS

	Columna 1	Columna 2	Columna 3	Valores de categorías/subcategorías
Caso 1				
Caso 2				
Caso 3				
Caso 4				
Caso k				

Los resultados del instrumento de medición se transfieren a la matriz por medio del libro de códigos. El proceso puede representarse así:



Si el libro de códigos no puede llevarse a cabo la transferencia. Vamos a suponer que hubiéramos aplicado la escala de actitud con tres ítems de la figura 9.21 a cuatro personas, obteniendo los siguientes resultados:

Persona 1

A continuación,
 1. "La Dirección de Impuestos Nacionales informa claramente sobre como, donde y cuándo pagar los impuestos"

El sujeto obtuvo:

(5) Muy de acuerdo (3) Ni de acuerdo, (2) En desacuerdo ni en desacuerdo	4 (de acuerdo)
<input checked="" type="checkbox"/> De acuerdo (1) Muy en desacuerdo	
2. "Los servicios que presta la Dirección de Impuestos Nacionales son en general muy buenos"	
<input checked="" type="checkbox"/> Muy de acuerdo (3) Ni de acuerdo, (2) En desacuerdo ni en desacuerdo	
(4) De acuerdo (1) Muy en desacuerdo	5 (muy de acuerdo)
3. "La Dirección de Impuestos Nacionales se caracteriza por la deshonestidad de sus funcionarios"	
(1) Muy de acuerdo <input checked="" type="checkbox"/> Ni de acuerdo, (4) En desacuerdo ni en desacuerdo	
(2) De acuerdo (5) Muy en desacuerdo	3 (ni de acuerdo, ni en desacuerdo)

Persona 2

Obtuvo respectivamente: 3 (ni de acuerdo, ni en desacuerdo)
 4 (de acuerdo)
 3 (ni de acuerdo, ni en desacuerdo)

Persona 3

Obtuvo respectivamente: 4
 4
 4

Persona 4

Obtuvo respectivamente: 5
 4
 3

De acuerdo con el libro de códigos (figura 9.21), tendríamos la siguiente matriz (figura 9.23):

EJEMPLO DE MATRIZ DE DATOS PARA EL LIBRO DE CÓDIGOS DE LA FIGURA 9.21

00196

	columna 1 (frase 1) (Infama)	columna 2 (frase 2) (servicios)	columna 3 (frase 3) (deshonestidad)	
Persona 1	4	5	3	← casos (en el ejemplo, sujetos)
Persona 2	3	4	3	
Persona 3	4	4	4	
Persona 4	5	4	3	

Valores de los sujetos en los ítems (en el ejemplo, frases) (categorías en las que cayeron transformadas a sus valores numéricos, es decir, codificadas)

En el ejemplo de Naves y Poplowsky (figura 9.22), la matriz sería la de la figura 9.24.

FIGURA 9.24

EJEMPLO HIPOTÉTICO DE MATRIZ DE DATOS PARA EL LIBRO DE CÓDIGOS DE LA FIGURA 9.22

	Columna 1 Número de sujeto	Columna 2 Número de sujeto	Columna 3 Tratamiento experimental	Columna 4 Distancia física	Columna 5 Movimientos corporales	Columna 6 Conducta visual	Columna 7 Conducta verbal	Columna 8 Codificador
S ₁	0	1	1	0	0	0	0	1
S ₂	0	2	1	0	0	0	0	1
S ₃	0	3	2	2	2	1	1	2
S ₄	0	4	1	1	0	0	0	2
S ₅	0	5	2	2	2	1	1	2
S ₆	0	6	1	0	1	1	0	3
S ₇	0	7	2	2	2	1	1	3
S ₈	0	8	2	2	2	1	1	3
S ₉	0	9	2	1	2	0	1	1
S ₁₀	1	0	1	2	1	0	0	2

El libro de códigos indica a los codificadores qué variable, ítem/categoría/sub-categoría va en cada columna y qué valores debe anotar en cada columna, así como el significado de cada valor numérico.

Con el libro de códigos sabemos que el sujeto 1 es el "01", que pertenece al "grupo cultural" ("1" en la tercer columna), que tuvo una conducta de alejamiento en su distancia física ("0" en la cuarta columna), que sus movimientos corporales fueron de tensión ("0" en la quinta columna), que su conducta visual fue a otra parte, no vio al sujeto ("0" en la sexta columna), que dijo frases dicótomas y/o silencios ("0" en la séptima columna) y que fue codificado por LRE ("1" en la octava columna). Y así con cada sujeto.

Obsérvese que, sin el libro de códigos, no se puede codificar y una matriz de datos carece de significado. La siguiente matriz no nos dice nada sin el libro de códigos. Por ejemplo:

5.	0	0	2	1	1	0	1	0	1	5
5	0	1	2	2	3	2	3	0	0	4
2	1	0	0	0	0	0	0	0	0	0
1	1	1	1	1	4	0	2	0	0	3
3	0	0	1	2	5	0	-2	0	0	3
3	1	0	1	2	5	0	1	0	1	3
4	0	1	0	2	4	0	1	0	1	4
5	1	1	2	2	1	0	0	0	1	4
2	0	1	2	2	2	0	2	1	0	5

¿Qué significa cada columna, cada dígito? Está en clave y sólo podemos tener acceso a ella mediante el libro de códigos.

En el libro de códigos y en la matriz de datos, una variable, ítem, categoría o subcategoría puede abarcar una, dos o más columnas, dependiendo de lo que esté indicando. Veamos el siguiente libro de códigos y la matriz correspondiente (figura 9.25).

FIGURA 9.25

EJEMPLO DE LA VARIABILIDAD QUE PUEDEN TENER LAS COLUMNAS

VARIABLE	LIBRO DE CÓDIGOS		
	CATEGORÍAS	CÓDIGOS	COLUMNA
— Número de sujeto	000 a 128	000 a 128	1, 2 / 3
— Sexo del respondiente	masculino femenino	1 2	4

— Edad del respondiente	00 a 77 No respondió	00 a 77 99	5 y 6 001
— Ingresos del respondiente	— Salario mínimo o menos — De dos a tres veces el salario mínimo — De cuatro a cinco veces el salario mínimo — De seis a ocho veces el salario mínimo — Nueve a once veces el salario mínimo — Doce a quince veces el salario mínimo — 16 o más veces el salario mínimo — No respondió	1 2 3 4 5 6 7 9	7
— Tamaño del lugar en que vive (m ²)	0000 a 9998 No respondió	0000 a 9998 9999	8,9,10 y 11

MATRIZ DE DATOS

VARIABLES

Columnas Sujetos	Número de sujeto			Sexo	Edad		Ingresos	Tamaño del lugar en que vive (residencia)			
	1	2	3	4	5	6	7	8	9	10	11

LA MATRIZ SE LLENA DE ACUERDO A LOS RESULTADOS

Valores perdidos

Cuando las personas no responden a un ítem o contestan incorrectamente o no puede registrarse la información (v.g., no se pudo observar la conducta), se crea una o varias categorías de valores perdidos y se les asignan sus respectivos códigos.

EJEMPLO

Si - 1	Si - 1
No - 2	No - 2
No contestó - 3	o Valor perdido por diversas razones
Contestó incorrectamente - 4	-9.

Hasta el momento se han presentado, por razones didácticas, ejemplos resumidos de libros de códigos. Desde luego, un libro de códigos normalmente tiene más variables o categorías y consecuentemente más columnas (al igual que la matriz de datos). Hay matrices que pueden tener 500 o más columnas. Asimismo, debe recordarse que los renglones son casos (sujetos, escuelas, series de televisión, etc.), y a veces es necesario extender las columnas a otro renglón u otros renglones (en computadoras limitadas a manejar 80 columnas). Esto podría representarse así:

S ₁	Col. 1	2	3	4	etc...	Col. 80
	Col. 81	82	83	84	etc...	Col. 160
	Col. 161	162	163	164	etc...	Col. 240
	Col. 241				etc...	Col. 320
S ₂	Col. 1	2	3	4		Col. 80
	Col. 81	82	83	84		Col. 160
	Col. 161	162	163	164		Col. 240
	Col. 241					Col. 320
S ₃						

Codificación física

El tercer paso del proceso de codificación es la codificación física de los datos, es decir, el llenado de la matriz de datos. Esta codificación la efectúan los codificadores, a quienes se les proporciona el libro de códigos. Así, cada codificador va vaciando las respuestas en la matriz de datos, de acuerdo con el libro de códigos. El vaciado de la matriz de datos puede hacerse en "hojas de tabulación", las cuales tienen columnas y renglones. En la figura 9.26 se muestra un ejemplo de una de estas hojas.

Si no alcanzan con una hoja de tabulación utilizan las hojas necesarias para vaciar los datos de todos los casos. Por ejemplo, la primera hoja puede alcanzarnos para 24 casos, pero si tenemos 200 casos, habremos de utilizar 9 hojas. Cada hoja estará llena de dígitos.

GRABADO Y GENERACIÓN DE ARCHIVOS

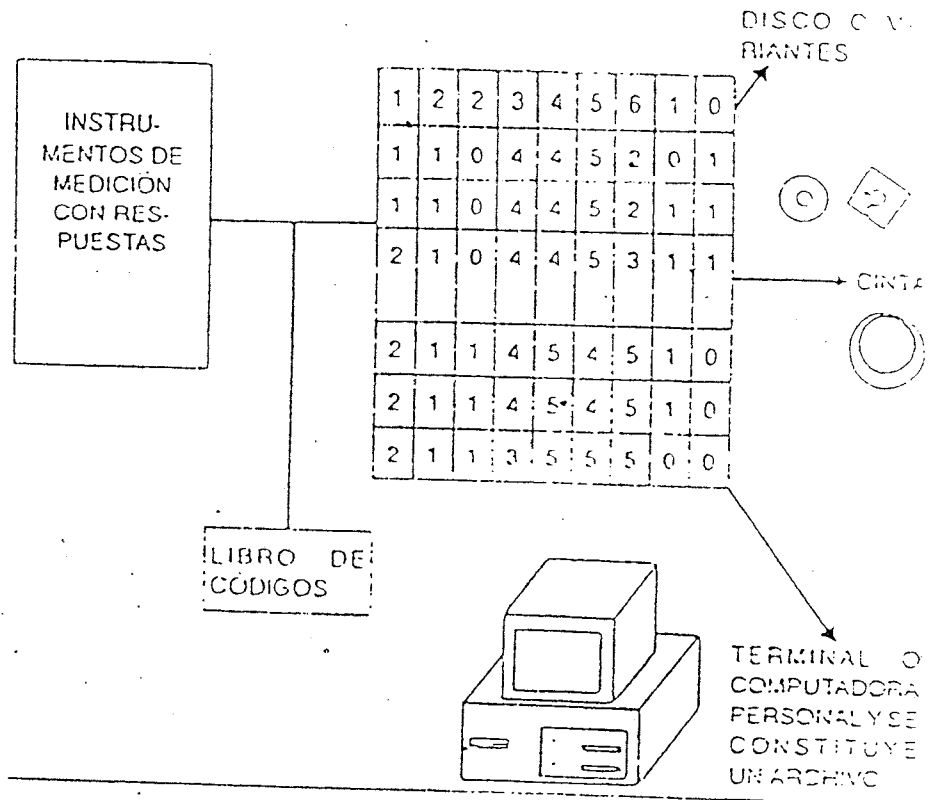
Las hojas de tabulación pueden copiarse a un disco o una cinta magnética para computadora, o bien, pueden telearse a un archivo dentro de una cuenta en una computadora. En cualquier caso, se crea un archivo o "file", el cual debe ser nombrado y contiene los datos codificados en valores numéricos en forma de matriz. El proceso se muestra en la figura 9.27.

También existen en la actualidad sistemas para la lectura óptica, los cuales pueden leer y almacenar los datos directamente de los cuestionarios (u otros instrumentos de medición) o de las hojas de tabulación. Desde luego, son sistemas costosos que requieren de lápices o tinta especiales. Asimismo, algunos investigadores con bastante experiencia pasan los datos directamente del instrumento aplicado al disco, cinta o computadora, pero se requiere de mucha práctica y personal capacitado.

Así, los datos han sido capturados en un archivo permanente y están listos para ser analizados mediante un programa de computadora. El proceso va desde la respuesta de los sujetos hasta un archivo que contiene una matriz (que es una matriz de valores numéricos que significan respuestas). La conducta y los valores de las variables han sido codificados.

FIGURA 9.27

PROCESO PARA CREAR UN ARCHIVO DE DATOS
(ARCHIVAR LA MATRIZ DE DATOS)



RESUMEN

1. Recolectar los datos implica seleccionar un instrumento de medición disponible o desarrollar uno propio, aplicar el instrumento de medición y preparar las mediciones obtenidas para que puedan analizarse correctamente.
2. Medir es el proceso de vincular conceptos abstractos con indicadores empíricos, mediante clasificación y/o cuantificación.
3. En toda investigación medimos las variables contenidas en las hipótesis.
4. Un instrumento de medición debe cubrir dos requisitos: confiabilidad y validez.
5. La confiabilidad se refiere al grado en que la aplicación repetida de un instrumento de medición al mismo sujeto u objeto, produce iguales resultados.
6. La validez se refiere al grado en que un instrumento de medición mide realmente la(s) variable(s) que pretende medir.
7. Se pueden aportar tres tipos de evidencia para la validez: evidencia relacionada con el contenido, evidencia relacionada con el criterio y evidencia relacionada con el constructo.

8. Los factores que principalmente pueden afectar la validez son: improvisación, utilizar instrumentos desarrollados en el extranjero y que no han sido validados a nuestro contexto, poca o nula empatía, factores de aplicación.
9. No hay medición perfecta, pero el error de medición debe reducirse a límites tolerables.
10. La confiabilidad se determina calculando un coeficiente de confiabilidad.
11. Los coeficientes de confiabilidad varían entre 0 y 1 (0 = nula confiabilidad, 1 = total confiabilidad).
12. Los procedimientos más comunes para calcular la confiabilidad son la medida de estabilidad, el método de formas alternas, el método de mitades parciales, el coeficiente alpha de Cronbach y el coeficiente K_T-20.
13. La validez de contenido se obtiene contrastando el universo de items contra los items presentes en el instrumento de medición.
14. La validez de criterio se obtiene comparando los resultados de aplicar el instrumento de medición contra los resultados de un criterio externo.
15. La validez de constructo se puede determinar mediante el análisis de factores.
16. Los pasos genéricos para construir un instrumento de medición son:
 - Listar las variables a medir
 - Revisar sus definiciones conceptuales y operacionales
 - Elegir una ya desarrollado o construir uno propio
 - Indicar niveles de medición de las variables (nominal, ordinal, por intervalos y de razón).
 - Indicar cómo se habrán de codificar los datos.
 - Aplicar prueba piloto.
 - Construir versión definitiva.
17. En la investigación social disponemos de diversos instrumentos de medición.
 - a) Principales escalas de actitudes: Likert, Diferencial Semántico y Guttman.
 - b) Cuestionarios (autoadministrado, por entrevista personal, por entrevista telefónica y por correo).
 - c) Análisis de contenido.
 - d) Observación.
 - e) Pruebas estandarizadas (procedimiento estándar)
 - f) Sesiones en profundidad.
 - g) Archivos y otras formas de medición.
18. Las respuestas se codifican.
19. La codificación implica:
 - a) Codificar los items o equivalentes no precodificados.
 - b) Elaborar el libro de códigos.
 - c) Efectuar físicamente la codificación.
 - d) Grabar y guardar los datos en un archivo permanente.

CONCEPTOS BÁSICOS

Recolección de datos
Medición
Instrumento de medición
Confiabilidad
Validez
Coeficiente de confiabilidad
Niveles de medición

- Medida de estabilidad
- Método de formas alternas
- Método de mitades paridas
- Coefficiente alfa de Cronbach
- Coefficiente KR-20 de Kuder-Richardson
- Evidencia relacionada con el contenido
- Evidencia relacionada con el criterio
- Evidencia relacionada con el constructo
- Escalas de actitudes
- Escala Likert
- Diferencial semántico
- Escalograma de Guttman
- Cuestionarios
- Análisis de contenido
- Observación
- Pruebas estandarizadas
- Pruebas proyectivas
- Sesiones en profundidad
- Codificación
- Codificador
- Hojas de codificación
- Matriz de datos
- Hojas de tabulación
- Archivo de datos

EJERCICIOS

1. Busque una investigación en algún artículo científico de una revista en ciencias sociales (ver apéndice uno) donde se incluya información sobre la confiabilidad y la validez del instrumento de medición. ¿El instrumento es confiable? ¿qué tan confiable? ¿qué técnica se utilizó para determinar la confiabilidad? ¿es válido? ¿cómo se determinó la validez?
2. Responda y discuta con ejemplos la diferencia entre confiabilidad y validez.
3. Defina ocho variables e indique su nivel de medición.
4. Suponga que alguien está tratando de evaluar la actitud hacia el Presidente de la República, construya un cuestionario tipo Likert con 10 ítems para medir dicha actitud e indique cómo se calificaría la escala total (5 ítems positivos y 5 negativos). Finalmente indique la dimensión que cada ítem pretende medir de dicha actitud (credibilidad, presencia física, etc.).
5. Construya un cuestionario para medir lo que usted considere conveniente (con preguntas demográficas y —por lo menos— 10 preguntas más), aplíquelo a 20 conocidos suyos, elabore el libro de códigos y la matriz de datos y vacíela en una hoja de tabulación elaborada por usted. Finalmente lea de la hoja de tabulación el significado de los dígitos de todas las columnas correspondientes a los 5 primeros casos.
6. Plantee una sesión en profundidad (indique objetivos, procedimiento, sujetos tipo, agenda, etc.) y organícela con amigos suyos. Al final, autoevalúe su experiencia.
7. Diseñe una investigación (planteamiento del problema, hipótesis, diseño) donde utilice por lo menos dos tipos de instrumentos de medición para recopilar los datos.

2. ¿Como se podrían aplicar el análisis de contenido y las sesiones en profundidad para la evaluación de un programa educativo a nivel superior?

BIBLIOGRAFÍA
SUGERIDA

- ANASTASI, A. (1992). *Psychological testing*. Nueva York, NY: MacMillan Publishing Co., Inc. Quinta Edición.
- BABBIE, E. R. (1979). *The practice of social research*. Belmont, CA: Wadsworth Publishing Co., Inc.
- CARMINES, E. G. y ZELLER, R. A. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage Publications, Inc. Series. *Quantitative Applications in the Social Sciences*, Vol. 17.
- CRONBACH, L. J. (1964). *Essentials of psychological testing*. Nueva York, NY: Gardner Press, Inc.
- KRIPPENDORFF, K. (1982). *Content analysis*. Beverly Hills, CA: Sage Publishing Co., Inc.
- MILLER D. C. (1977). *Handbook of research design and social measurement*. Nueva York, NY: Longman, Inc. Tercera edición.
- NUNNALLY, J. C. (1978). *Introduction to psychological measurement*. Nueva York, NY: McGraw-Hill, Inc.
- THORNDIKE, R. L. y HAGEN, E. (1980). *Tests y técnicas de medición en psicología y educación*. México, D.F.: Editorial Trillas.
- WEBB, E. J.; CAMPBELL, D. T., y SCHWARTZ, R.D. (1966). *Unobtrusive measures nonreactive research in the social sciences*. USA: Rand McNally College Publishing Co.
- WIEBMAN, W. (1986). *Research: methods in education*. Newton, Mass: Allyn and Bacon, Inc. Cuarta edición, capítulo 11 ("Measurement and data collection").

EJEMPLO

La televisión y el niño

FIGURA 9.2ª

QUESTIONARIOS QUE SE ADMINISTRARON A LOS NIÑOS DEL DISTRITO FEDERAL PARA EL EJEMPLO DE LA TELEVISIÓN Y EL NIÑO (SE INCLUYE PARCIALMENTE EN VERSIÓN PRECODIFICADA)

1. ¿Cada cuánto lees un periódico?
- _____ casi nunca (0)
- _____ una vez al mes (1)
- _____ una vez cada semana (2)
- _____ 2 o 3 veces a la semana (3)
- _____ 4 o 5 veces a la semana (4)
- _____ diario (5)
2. ¿Cada cuánto vas al cine?
- _____ una vez por semana (1)
- _____ una cada dos semanas (2)
- _____ una vez al mes (3)
- _____ una vez cada 2 o 3 meses (4)
- _____ casi nunca (0)

Me gusta mucho
(2)

Me gusta un poco
(1)

No lo veo
(0)

005
00206

- Fútbol
- Películas
- Los Picapietra
- Hechizada
- Ven conmigo
- La criada bien criada
- Kojack
- En punto
- Plaza Sésamo
- El Oso Ruperto
- Locos Adams
- Una muchacha llamada Milagros
- Los polivoces
- Viaje al fondo del mar
- 24 horas
- Universo 5
- Clásicos infantiles
- Nany y el profesor
- Barata de primavera
- El chavo del 8
- El llanero solitario
- Deporteve
- Platicame un libro

18. ¿Generalmente con quién ves la televisión?

- _____ solo (1)
- _____ con mi papá (2)
- _____ con mi mamá (3)
- _____ con mis hermanos (4)
- _____ con mis primos (5)
- _____ con mis amigos (6)
- _____ con la sirvienta (7)

2. ¿Como se podrían aplicar el análisis de contenido y las secciones en profundidad para la evaluación de un programa educativo a nivel superior?

BIBLIOGRAFÍA SUGERIDA

- ANASTASI, A. (1982). *Psychological testing*. Nueva York, NY: MacMillan Publishing Co., Inc. Quinta Edición.
- BABBIE, E. R. (1979). *The practice of social research*. Belmont, CA: Wadsworth Publishing Co., Inc.
- CARMINES, E. G. y ZELLER, R. A. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage Publications, Inc. Series. Quantitative Applications in the Social Sciences, Vol. 17.
- CRONBACH, L. J. (1964). *Essentials of psychological testing*. Nueva York, NY: Gardner Press, Inc.
- KRIPPENDORFF, K. (1982). *Content analysis*. Beverly Hills, CA: Sage Publishing Co., Inc.
- MILLER D. C. (1977). *Handbook of research design and social measurement*. Nueva York, NY: Longman, Inc. Tercera edición.
- NUNNALLY, J. C. (1978). *Introduction to psychological measurement*. Nueva York, NY: McGraw-Hill, Inc.
- THORNDIKE, R. L. y HAGEN, E. (1980). *Tests y técnicas de medición en psicología y educación*. México, D.F.: Editorial Trillas.
- WEBB, E. J.; CAMPBELL, D. T. y SCHWARTZ, R.D. (1966). *Unobtrusive measures: nonreactive research in the social sciences*. USA: Rand McNally College Publishing Co.
- WIERSSMA, W. (1986). *Research methods in education*. Newton, Mass: Allyn and Bacon, Inc. Cuarta edición, capítulo 11 ("Measurement and data collection").

EJEMPLO

La televisión y el niño

FIGURA 9.2ª

QUESTIONARIOS QUE SE ADMINISTRARON A LOS NIÑOS DEL DISTRITO FEDERAL PARA EL EJEMPLO DE LA TELEVISIÓN Y EL NIÑO (SE INCLUYE PARCIALMENTE EN VERSIÓN PRECODIFICADA)

1. ¿Cada cuánto lees un periódico?
 - _____ casi nunca (0)
 - _____ una vez al mes (1)
 - _____ una vez cada semana (2)
 - _____ 2 o 3 veces a la semana (3)
 - _____ 4 o 5 veces a la semana (4)
 - _____ diario (5)

2. ¿Cada cuánto vas al cine?
 - _____ una vez por semana (1)
 - _____ una cada dos semanas (2)
 - _____ una vez al mes (3)
 - _____ una vez cada 2 o 3 meses (4)
 - _____ casi nunca (0)

Me gusta mucho

(2)

Me gusta un poco

(1)

No lo veo

(0)

000

00206

- _____ Fútbol
- _____ Películas
- _____ Los Picapiedra
- _____ Hechizada
- _____ Ven conmigo
- _____ La criada bien criada
- _____ Kojack
- _____ En punto
- _____ Plaza Sésamo
- _____ El Oso Ruperto
- _____ Locos Adams
- _____ Una muchacha llamada Milagros
- _____ Los polivoces
- _____ Viaje al fondo del mar
- _____ 24 horas
- _____ Universo 5
- _____ Clásicos infantiles
- _____ Nany y el profesor
- _____ Barata de primavera
- _____ El chavo del 8
- _____ El llanero solitario
- _____ Deporteve
- _____ Platicame un libro

18. ¿Generalmente con quién ves la televisión?

- _____ solo (1)
- _____ con mi papá (2)
- _____ con mi mamá (3)
- _____ con mis hermanos (4)
- _____ con mis primos (5)
- _____ con mis amigos (6)
- _____ con la sirvienta (7)

19. La mayoría de las veces, ¿quién escoge los programas?
- _____ yo (1)
 _____ papá (2)
 _____ mamá (3)
 _____ hermanos (4)
 _____ primos (5)
 _____ amigos (6)
 _____ otra (7)
20. Marca con una palomita, si estas de acuerdo o no con las siguientes frases.
 Mi papá o mi mamá me regañan cuando veo mucha televisión.
- Si _____ No _____
 (1) (0)
- A veces me castigan sin ver televisión.
- Si _____ No _____
 (1) (0)
- Mi papá o mi mamá me prohiben ver algunos programas de televisión.
- Si _____ No _____
 (1) (0)
- Yo me voy a dormir a la hora que quiero.
- Si _____ No _____
 (1) (0)
21. Veo la televisión porque me río mucho.
- _____ siempre (2)
 _____ a veces (1)
 _____ casi nunca (0)
22. Veo la televisión porque es muy divertida.
- _____ siempre (2)
 _____ a veces (1)
 _____ casi nunca (0)
23. Veo la televisión porque se me pasa el tiempo volando.
- _____ siempre (2)
 _____ a veces (1)
 _____ casi nunca (0)
24. Veo la televisión porque me entretiene mucho.
- _____ siempre (2)
 _____ a veces (1)
 _____ casi nunca (0)

25. ¿Ves la televisión porque nunca me aburre?

_____ siempre (2)

_____ a veces (1)

_____ casi nunca (0)

26. Cuando ves la televisión por las razones de arriba ¿qué tanto te gusta?

_____ muchísimo (4)

_____ mucho (3)

_____ regular (2)

_____ un poco (1)

_____ casi nada (0)

IV BIBLIOGRAFIA GENERAL

- CEMBRANOS, F; D. H. Montesinos y Maria Bustelo La animación sociocultural: propuesta metodológica. Madrid, Editorial popular, 1989
- DEVEREUX, George. De la ansiedad al método en las ciencias del comportamiento. México, siglo XXI 1991.
- GALINDO C. Como coordinar el nivel preescolar y el ciclo inicial. Madrid Edit. Escuela Española, 1984.
- HERNANDEZ R. y Cols. Metodología de la investigación, México, Mcgraw-Hill 1990
- KARMEI, Louis; Medición y evaluación escolar; México, Trillas 1986.
- S. E. C. "Estrategias de evaluación en el aula" Boletín del CETE No. 9-10 Veracruz, julio-diciembre de 1994.
- WHEELER El desarrollo del currículum escolar, España, Santillana-aula XXI, 1985.

00210

UNIVERSIDAD PEDAGOGICA NACIONAL
AREA DE SERVICIOS DE BIBLIOTECA
Y DE APOYO ACADEMICO
FECHA DE DEVOLUCION

API

4-IX-10

El lector se obliga a devolver este material antes
del vencimiento del préstamo señalado por el último sello

SILVIA SCHMELKES
ASESORA EXTERNA

COORDINACION DEL PROYECTO LE '94
XOCHITL L. MORENO FERNANDEZ

JULIO DE 1995.