

Capítulo 9 Recolección de los datos

OBJETIVOS DE APRENDIZAJE

- Al terminar este capítulo, el alumno será capaz de:
- Comprender el significado de "medir" en ciencias sociales.
 - Comprender los requisitos que toda medición debe cumplir: confiabilidad y validez.
 - Reconocer los principales instrumentos de medición disponibles en ciencias sociales.
 - Elaborar y aplicar diferentes instrumentos de medición.
 - Preparar datos para su análisis.

PROCESO DE INVESTIGACION

Ocaso, 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 su análisis.

SINTEISIS

El capítulo presenta una definición de medición en el contexto de las ciencias sociales, así como los requisitos que todo instrumento de medición debe reunir: confiabilidad y validez. Diversos métodos para determinar la confiabilidad y validez son revisados.

Además, el capítulo analiza y ejemplifica las principales maneras de medir en ciencias sociales: escalas de actitudes, cuestionarios, análisis de contenido, observación, pruebas estandarizadas, sesiones en profundidad y utilización de archivos.

Finalmente, en el capítulo se presenta el procedimiento de codificación de los datos obtenidos y la forma de prepararlos para el análisis.

9.1. ¿QUÉ IMPLICA LA ETAPA DE RECOLECCIÓN DE LOS DATOS?

Una vez que seleccionamos el diseño de investigación apropiado y la muestra adecuada de acuerdo con nuestro problema de estudio e hipótesis, la siguiente etapa consiste en recolectar los datos pertinentes sobre las variables involucradas en la investigación.

Recolectar los datos implica tres actividades estrechamente vinculadas entre sí:

- a) *Selección un instrumento de medición de los disponibles en el estudio del comportamiento o desarrollar uno (el instrumento de recolección de los datos).* Este instrumento debe ser válido y confiable, de lo contrario no podemos basarnos en sus resultados.
- b) *Aplicar ese instrumento de medición.* Es decir, obtener las observaciones y mediciones de las variables que son de interés para nuestro estudio (medir variables).
- c) *Preparar las mediciones obtenidas para que puedan analizarse correctamente (a esta actividad se le denomina codificación de los datos).*

9.2. ¿QUÉ SIGNIFICA MEDIR?

De acuerdo con la definición clásica del término, ampliamente difundida, *medir* significa "asignar números a objetos y eventos de acuerdo con reglas" (Stevens, 1951). Sin embargo, como señalan Carnines y Zeller (1979), esta definición es más apropiada para las ciencias físicas que para las ciencias sociales, ya que varios de los fenómenos que son medidos en éstas no pueden caracterizarse como objetos o eventos, puesto que son demasiado abstractos para ello. La distinción cognitiva, la alienación, el producto nacional bruto y la credibilidad son conceptos tan abstractos para ser considerados "cosas que pueden verse a tocarse" (definición de objeto) o solamente como "resultado, consecuencia o producto" (definición de evento) (Carnines y Zeller, 1979, p. 10).

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 (Carnines 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 (Carnines y Zeller, 1979). Así, los registros del instrumento de medición representan valores observables de conceptos abstractos. *Un instrumento de medición ade-*

cuado es aquel que registra datos observables que representan verdaderamente 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 es un hecho que debemos acercarnos lo más posible a la representación fiel de las variables a observar, mediante el instrumento de medición que desarrollemos.

9.3. ¿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 se midiera en este momento la temperatura ambiental mediante un termómetro e indicara que hay 22°C. Un minuto más tarde se consultara otra vez el termómetro y éste indicara que hay 5°C. Tres minutos después se observara el termómetro y éste indicara que hay 40°C. Ese termómetro no sería confiable (su aplicación repetida produce resultados distintos). Igualmente, si una prueba de inteligencia se aplica hoy a un grupo de personas y proporciona ciertos valores de inteligencia; se aplica un mes después y proporciona valores diferentes, al igual que en subsecuentes mediciones, tal prueba no es confiable (analicémos 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 9.1

Ejemplo de resultados proporcionados por un instrumento de medición sin confiabilidad.

	Primera aplicación	Segunda aplicación	Tercera aplicación
Martha	130	Laura	131
Laura	125	Luis	130
Arturo	118	Marco	127
Luis	112	Arturo	120
Marco	110	Chester	118
Rosa María	110	Teresa	118
Chester	108	Martha	115
Teresa	107	Rosa María	107
		Luis	140
		Teresa	129
		Martha	124
		Rosa María	120
		Laura	109
		Chester	108
		Arturo	103
		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 muchos otros 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: ¿Esta midiendo lo que 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

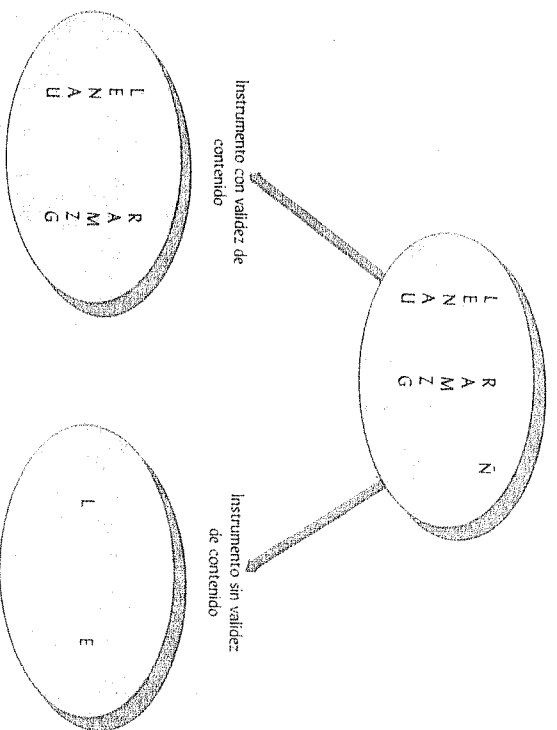
La *validez de contenido* se refiere al grado en que un instrumento refleja un dominio específico de contenido de lo que se mide. Es el grado en que la medición representa al concepto medido (Bohrnsstedt, 1976). Por ejemplo, una prueba de operaciones aritméticas no tendrá validez de contenido si incluye solo problemas de resta y excluye problemas de suma, multiplicación o división (Carmine y Zeller, 1979). O bien, una prueba de conocimientos sobre las canciones de *The 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 comparándolo con algún criterio externo. Este criterio es un estándar con el que se juzga la validez del instrumento (Wiersma, 1986). Entre más se relacionen los resultados del instrumento de medición con el criterio, la validez del instrumento 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 aeroplano.

FIGURA 9.2 Ilustración de un instrumento de medición con validez de contenido versus uno que carece de ella.



Si el criterio se fija en el *presente*, se habla de *validez concurrente* (los resultados del instrumento se correlacionan con el criterio en el mismo momento o punto del tiempo). Por ejemplo, un cuestionario para detectar las preferencias del electorado 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 el criterio se fija en el *futuro*, se habla de *validez predictiva*. Por ejemplo, una prueba para determinar la capacidad administrativa de otros ejecutivos se puede validar comparando sus resultados con el futuro desempeño de los ejecutivos medidos.

3. Evidencia relacionada con el constructo

La *validez de constructo* es probablemente "la más importante sobre todo desde una perspectiva científica y se refiere al grado en que una medición se relaciona

¹⁵ La explicación se basa en Carmine y Zeller (1979).

consistentemente con otras mediciones de acuerdo con hipótesis derivadas teóricamente y que concierne 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, suponamos que un investigador desea evaluar la *validez de constructo* de una medición particular, digamos una escala de motivación intrínseca: "el Cuestionario de Reacción a Tareas", versión mexicana (Hernández-Sampieri y Cortes, 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 validez de constructo del Cuestionario de Reacción a Tareas, versión mexicana (la validez para medir la motivación intrínseca).

La *validez 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 con el nivel en que clarifica la validez 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 una teoría sumamente desarrollada, pero sí 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.

Para analizar las posibles interpretaciones de evidencia negativa en la *validez de constructo*, se sugiere consultar a Cronbach y Mechi (1955) y Cronbach (1984).

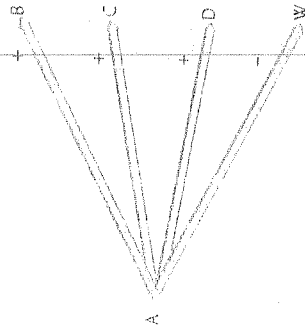
VALIDEZ DE CONTENIDO, VALIDEZ DE CONSTRUCCIÓN Y VALIDEZ DE CONFIABILIDAD

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

FIGURA 9.3
Presentación
gráfica de un
instrumento con
validez de
constructo.

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".

criterio y validez de constructo tenga un instrumento de medición; éste 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 pueden tomarse 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 aun en ambientes académicos).

Aun a los investigadores experimentados les toma cierto tiempo desarrollar un instrumento de medición. Es por ello que los construyen con cuidado y frecuentemente están desarrollándolos, para que cuando los necesitan con premura se encuentren preparados para aplicarlos, pero no los improvisan. Además, para poder construir un instrumento de medición se requiere conocer muy bien 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 ni remotamente validarlo. Es un primer y necesario paso, pero solo 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 "arraigado". Las culturas, los grupos y las personas cambian; y esto debemos tomarlo en cuenta al elegir o desarrollar un instrumento de medición.

Otro factor es que en ocasiones el instrumento resulta inadecuado para las personas a las que se les aplicará no es empático. Utilizar un lenguaje muy elevado para el encuestado, 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 entrevistados, son errores que pueden afectar la validez y confiabilidad del instrumento de medición.

El cuarto factor que puede influir está constituido por las condiciones en las que se aplica el instrumento de medición. El ruido, el frío (por ejemplo en una encuesta de casa en casa), un instrumento demasiado largo o tedioso, son condiciones 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 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 ¿COMO SE SABE SI UN INSTRUMENTO DE MEDICIÓN ES CONFIABLE Y VALIDO?

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 (en el cual 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.

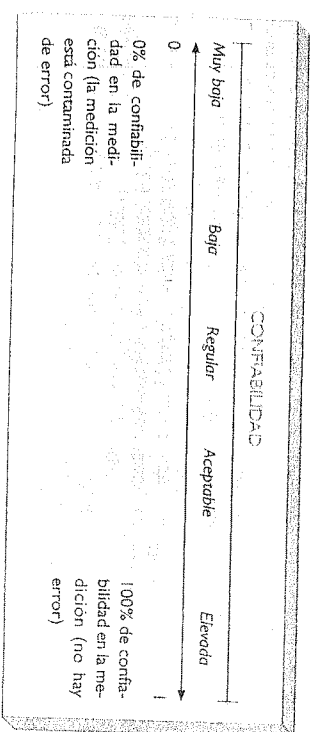
Cálculo de la confiabilidad

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 más se acerque el coeficiente a cero (0), hay mayor error en la medición. Esto se ilustra en la figura 9.4.

Los procedimientos más utilizados para determinar la confiabilidad median un coeficiente son:

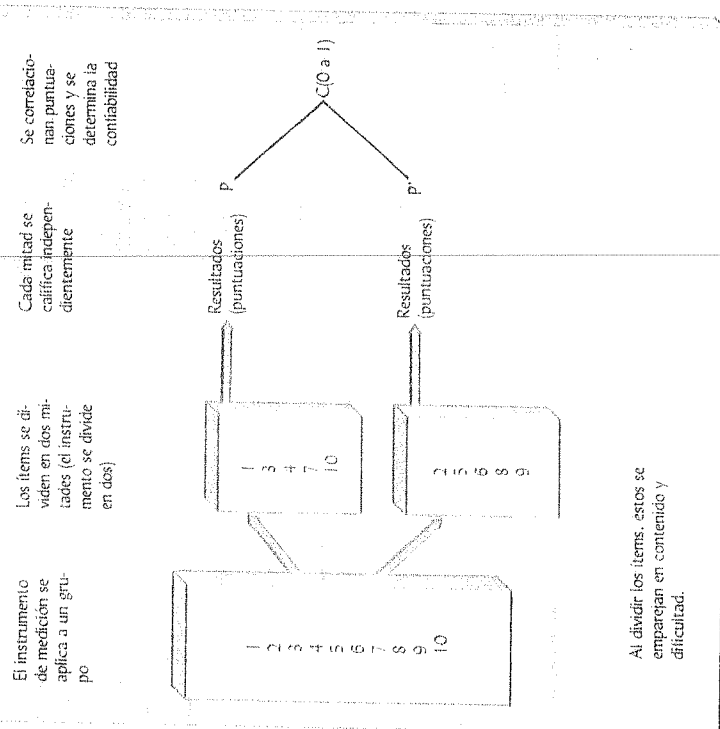
Método de estabilidad (confiabilidad por test-retest). En este procedimiento un mismo instrumento de medición (o ítems o indicadores)¹⁶ es aplicado dos o más veces a un mismo grupo de personas, después de cierto periodo. Si la

FIGURA 9.4
Interpretación de un coeficiente de confiabilidad



¹⁶ Un ítem es la unidad mínima que compone a una medición; es un reactivo que sustituye una respuesta en un sujeto (por ejemplo, una pregunta, una frase, una lamina, fotografía, un objeto de descripción).

FIGURA 9.5
Esquema del
procedimiento de
mitades-partidas.



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 periodo de tiempo entre las mediciones es un factor a considerar. Si el periodo es largo y la variable susceptible de cambios, ello puede confundir la interpretación del coeficiente de confiabilidad obtenido por este procedimiento. Y si el periodo es corto las personas pueden recordar como procedieron en la primera aplicación del instrumento, para aparecer como más consistentes de lo que son en realidad (Bohrnstedt, 1976).

1. *Método de formas alternativas o paralelas.* En este procedimiento no se administra el mismo instrumento de medición, sino dos o más versiones equivalentes de este. 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 periodo 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 mitades partidas (split-halves).* Los procedimientos anteriores (medida de estabilidad y método de formas alternativas), requieren cuando menos dos administraciones de la medición en el mismo grupo de individuos. En cambio, el método de mitades-partidas requiere solo 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 mostrar también una baja puntuación en la otra mitad. El procedimiento de diagrama en la figura 9.5.

La confiabilidad varía de acuerdo con el número de ítems que incluya el instrumento de medición. Cuantos más ítems mayor es la confiabilidad. Esto resulta lógico, veámoslo con un ejemplo cotidiano: Si se desea probar que tan confiable o consistente es la lealtad de un amigo hacia nuestra persona, cuantas más pruebas le pongamos, su confiabilidad será mayor. Claro esta que demasiados ítems provocarán cansancio en el respondiente.

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 explica en el siguiente capítulo.

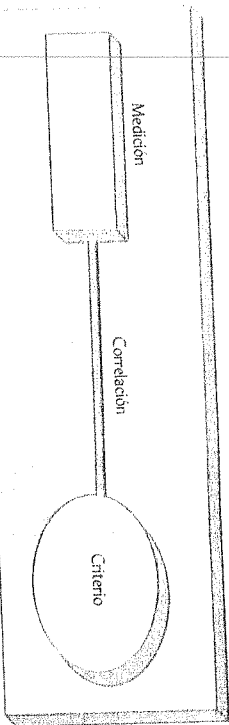
3. *Coefficiente KR-20.* Kuder y Richardson (1937) desarrollaron un coeficiente para estimar la confiabilidad de una medición, su interpretación es la misma que la del coeficiente alfa.

Cálculo de la validez

Resulta complejo obtener la validez de contenido. Primero, es necesario revisar cómo ha sido utilizada la variable por otros investigadores. Y con base en dicha

revisión elaborar un universo de ítems posibles para medir la variable y sus dimensiones (el universo debe 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 está compuesta por diversas dimensiones o facetas, 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) (Bohrnstedt, 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 se lo



ma como coeficiente de validez (Bohrnstedt, 1976). Esto podría representarse así:

La validez de constructo suele determinarse mediante un procedimiento denominado "análisis de factores". Su aplicación requiere sólidos conocimientos estadísticos y un programa estadístico apropiado de computadora. Para quien desee competirse con esta técnica recomendamos consultar Hartman (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 Curtis (1976) y, en español, Padua (1979). Aunque es requisito conocer el programa estadístico para computadora. Esta técnica se describe en el apartado 10.10.4.

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.

Pases

1. Listar las variables que se pretenden medir u observar.
2. Revisar su definición conceptual y comprender su significado. Por ejemplo, comprender bien que es la motivación intrínseca y qué dimensiones la integran.

Ver Véase el tema de correlación en el siguiente capítulo.

Revisar cómo han sido definidas operacionalmente las variables, esto es, como 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).

A) Elegir el instrumento o los instrumentos (ya desarrollados) que hayan sido favorecidos por la comparación y adaptarlos al contexto de la investigación. Para este caso sólo deben seleccionarse instrumentos cuya confiabilidad y validez se reporte. No se puede confiar en una forma 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. Recuérdese que la primera varía 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). Tampoco no debe olvidarse que traducir no es validar un instrumento, por muy buena que sea la traducción.

FIGURA 9.6
Ejemplo de desarrollo de ítems.
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 previos e ítems para cada dimensión. La figura 9.6 es un ejemplo de ello:

Variable	Definición operacional	Indicadores	Dimensiones	Ítems
Coordinación entre organizaciones compradoras y proveedores, desde el punto de vista de las primeras.	Grado percibido de esfuerzo invertido para no provocar problemas a la otra parte al interferir en sus deberes y responsabilidades.	Grado percibido de buena voluntad de ambas partes.	Coordinación de conflictos.	¿Cuanto se esfuerza su empresa por no provocar problemas con sus proveedores? 1. Se esfuerza al máximo posible. 2. Se esfuerza poco. 3. Se esfuerza medianamente. 4. Se esfuerza mucho. 5. Se esfuerza al máximo posible. ¿Cuánto se esfuerza su proveedor por no provocar problemas con su empresa?

FIGURA 9.6 (Continuación).

Variable	Definición operacional	Indicadores	Dimensiones	Ítems
				<ol style="list-style-type: none"> 1. Se esfuerzan al mínimo posible. 2. Se esfuerzan poco. 3. Se esfuerzan medianamente. 4. Se esfuerzan mucho. 5. Se esfuerzan al máximo posible. <p>¿Cuanto se esfuerza su empresa por no interferir en los deberes y responsabilidades de sus proveedores?</p> <ol style="list-style-type: none"> 5. Se esfuerza al máximo posible. 4. Se esfuerza mucho. 3. Se esfuerza medianamente. 2. Se esfuerza poco. 1. Se esfuerza al mínimo posible. <p>¿Cuanto se esfuerzan sus proveedores por no interferir con los deberes y responsabilidades de su empresa?</p> <ol style="list-style-type: none"> 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 no interferencia.	
				Coordinación de objetivos.

FIGURA 9.6 (Continuación).

Variable	Definición operacional	Indicadores	Dimensiones	Ítems
				<p>para alcanzar objetivos comunes?</p> <ol style="list-style-type: none"> 5. Se esfuerza al máximo posible. 4. Se esfuerza mucho. 3. Se esfuerza medianamente. 2. Se esfuerza poco. 1. Se esfuerza al mínimo posible. <p>¿Cuanto se esfuerzan los proveedores por trabajar junto con su empresa, de manera constante, para alcanzar objetivos comunes?</p> <ol style="list-style-type: none"> 5. Se esfuerzan al máximo posible. 4. Se esfuerzan mucho. 3. Se esfuerzan medianamente. 2. Se esfuerzan poco. 1. Se esfuerzan al mínimo posible. <p>En general, ¿qué tan bien establecidas están las rutinas para el trato de la empresa con sus proveedores?</p> <ol style="list-style-type: none"> 5. Muy bien establecidas. 4. Bien establecidas. 3. Medianamente establecidas. 2. Mal establecidas. 1. Muy mal establecidas.
			Coordinación de objetivos.	
				Coordinación de rutinas.
				<p>Estableciendo un promedio aproximado ¿con qué frecuencia recibe su</p>

Frecuencia de la interacción entre organizaciones.

Lapsos de interacciones entre organizaciones.

Visitas de representantes.

FIGURA 9.2 (Continuación)

Variable	Definición operacional	Indicadores	Dimensiones	Ítems
				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. 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 ¿con qué frecuencia 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.

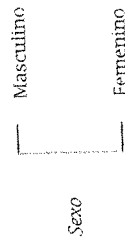
FIGURA 9.3 (Continuación)

Variable	Definición operacional	Indicadores	Dimensiones	Ítems
			Llamadas telefónicas.	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). Estableciendo un promedio aproximado ¿con qué frecuencia le llaman por teléfono a su empresa los representantes de sus proveedores más 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). Etcétera.

En este segundo caso, debemos asegurarnos de tener un número suficiente 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.

- 2) **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. 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.

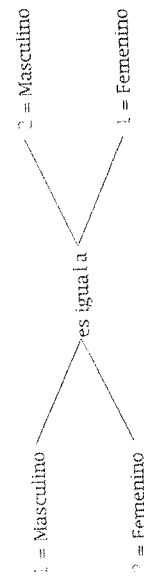


Si les asignamos una etiqueta o símbolo a cada categoría, esto identifica exclusivamente a la categoría. Por ejemplo:

x = Masculino

z = Femenino

Si usamos numerales es lo mismo:



Los números utilizados en este nivel de medición tienen una función puramente de clasificación 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

$$1 + 2 = 3$$

- 3 = Protestante
- 4 = Musulmán
- 5 = Otros

Un católico + un judío = protestante?
(no tiene sentido)

Las variables nominales pueden incluir dos categorías (dicotómicas), o bien, tres o más categorías (categóricas). Ejemplos de variables nominales dicotómicas sería el sexo y el tipo de escuela a la que se asiste (privada-pública); y de nominales categóricas tendríamos a la afiliación política (Partido A, Partido B,...), la carrera elegida, la raza, el departamento o provincia o estado de nacimiento y el canal de televisión preferido.

- 2) **Nivel de medición ordinal.** En este nivel hay 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 Estados Unidos ha sido medido por diversas escalas que reordenan a las profesiones de acuerdo con su prestigio, por ejemplo:²⁶

Valor en la escala

Profesión

90
80
60
50
02

Ingeniero químico
Científico de ciencias naturales (excluyendo la química)
Actor
Operador de estaciones eléctricas de potencia
Manufactureros de tabaco

90 es más que 80, 80 más que 60, 60 más que 50 y así sucesivamente; es decir, 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 ciencias naturales (80) y un ingeniero químico (90). Aparentemente en ambos casos la distancia es 10, pero no es una distancia real. Otra escala²⁷ clasificó el prestigio de dichas profesiones de la siguiente manera:

Valor en la escala

Profesión

98
95
84
78
13

Ingeniero químico
Científico de ciencias naturales (excluyendo la química)
Actor
Operador de estaciones eléctricas de potencia
Manufactureros de tabaco

²⁶ Duncan, 1961.

²⁷ Nam et al. (1975).

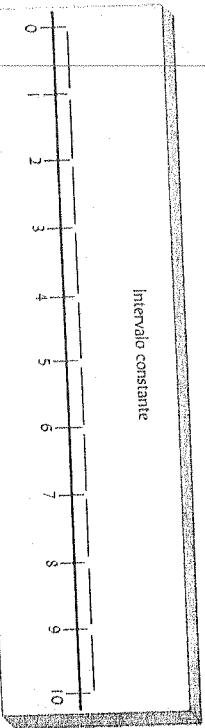
Aquí la distancia entre un actor (84) y un operador de estaciones (78) es de 6, y la distancia entre un ingeniero químico (98) y un científico de ciencias naturales (95) es de 3.

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
Intendencia	1

Sabemos que el presidente (10) es más que el vicepresidente (9), este 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 (Empleado A) + 5 (jefe) = 9 (vicepresidente) ni que 10 (presidente) + 5 (jefe) = 2 (Empleado C). Sería absurdo, no tiene sentido.

3. *Nivel de medición por intervalos.* Además del 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 un intervalo constante, una unidad de medida.

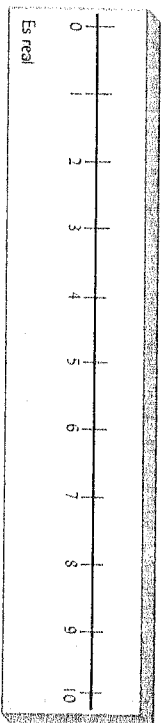


Por ejemplo: Una prueba de resolución de problemas matemáticos (30) produce la misma distancia de igual dificultad. Si Ana Cecilia resolvió 10, Laura resolvió 20 y Brenda 30. La distancia entre Ana, Cecilia y Laura es igual a la distancia entre Laura y Brenda.

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 (ninguna) 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 y algunas estadísticas modernas, que de otro modo no se usarían. Aunque algunos investigadores no están de acuerdo en suponer tales mediciones como si fueran de intervalo.

4. *Nivel de medición de razón.* 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 sería la exposición a la televisión, el número de hijos, la productividad, las ventas de un producto y el ingreso.

Desde luego, hay variables que puedan 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 K días)
— Ordinal	Bastante antigüedad
	Antigüedad regular
	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.

7. *Indicar cómo 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 fuéramos la variable "sexo" con sus respectivas categorías, "masculino" y "femenino", a cada categoría le asignaríamos un valor. Este podría ser:

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 Larios un "2" y así sucesivamente.

Otro ejemplo sería la variable "horas de exposición diaria a la televisión", que podría codificarse de la siguiente manera:

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?" Las categorías podrían ser muchas más de las que nos imaginemos y resultaría difícil predecir con precisión cuantas y cuales serán. En estos casos la codificación se lleva a cabo una vez que se aplica el ítem (*a posteriori*).

La codificación es necesaria para analizar cuantitativamente los datos (aplicar análisis estadístico). A veces se utilizan letras o símbolos en lugar de números (*, A, Z).

9) 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.

10) 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 adaptarse a las características de los tipos de instrumentos de que disponemos en el estudio del comportamiento, que estudiaremos 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 pueden ser combinados varios métodos de recolección de los datos. A continuación los describimos brevemente.

9.6.1 Escalas para medir las actitudes

Una *actitud* es una predisposición aprendida para responder consistentemente de una manera favorable o desfavorable ante un objeto de 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, 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. Por ello las mediciones de actitudes deben interpretarse como "síntomas" y no como "hechos" (Padua, 1979). Si detecto que la actitud de un grupo hacia la contaminación es desfavorable, esto no significa que las personas estén 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 Likert¹⁶

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. Es decir, se presenta cada afirmación y se pide al sujeto que extirpe 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 20 palabras.

Ejemplo	Afirmación
<i>Objeto de actitud medido</i>	<i>Afirmación</i>
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 cuanto 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.

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

¹⁶ Para profundizar en esta técnica se recomienda consultar a Likert (1976a o 1976b), Sailer y Hough (1976) y Padua (1979).

() Ni de acuerdo, ni en desacuerdo
 () En desacuerdo
 () Muy en desacuerdo

O bien utilizando recuadros en lugar de parentesis:

Definitivamente sí
 Probablemente sí
 Indeciso
 Probablemente no
 Definitivamente no

FIGURA 9.7
 Alternativas o puntos en las escalas Likert.

Es indispensable señalar que el número de categorías de respuesta debe ser el mismo para todas las afirmaciones.

Alternativa 1:	"Afirmación"
Muy de acuerdo	Ni de acuerdo, ni en desacuerdo
De acuerdo	En desacuerdo
Totallymente de acuerdo	Muy en desacuerdo
Alternativa 2:	"Afirmación"
De acuerdo	Neutral
De acuerdo	En desacuerdo
De acuerdo	Totallymente en desacuerdo
Alternativa 3:	"Afirmación"
Definitivamente sí	Indeciso
Probablemente sí	Probablemente no
Indeciso	Definitivamente no
Alternativa 4:	"Afirmación"
Completamente verdadero	Falso
Verdadero	Falso
Completamente verdadero	Completamente falso

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 cuanto los sujetos estén más de acuerdo con la afirmación, su actitud será 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 que 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 9.3
Ejemplo de una
escala Likert

Las afirmaciones que voy a leer 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.

1. El personal de la Dirección General de Impuestos Nacionales es grosero al atender al público.
 - 1. Muy de acuerdo
 - 2. De acuerdo
 - 3. Ni de acuerdo, ni en desacuerdo
 - 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
 - 2. De acuerdo
 - 3. Ni de acuerdo, ni en desacuerdo
 - 4. En desacuerdo
 - 5. Muy en desacuerdo
3. Los servicios que presta la Dirección General de Impuestos Nacionales son en general muy buenos.
 - 1. Muy de acuerdo
 - 2. De acuerdo
 - 3. Ni de acuerdo, ni en desacuerdo
 - 4. En desacuerdo
 - 5. Muy en desacuerdo
4. La Dirección General de Impuestos Nacionales informa claramente sobre cómo, dónde y cuándo pagar los impuestos.
 - 1. Muy de acuerdo
 - 2. De acuerdo
 - 3. Ni de acuerdo, ni en desacuerdo
 - 4. En desacuerdo
 - 5. Muy en desacuerdo

⁴¹ El ejemplo fue utilizado en un país latinoamericano y su confiabilidad total fue de 0.83; aquí se presenta una versión reducida de la escala original. El nombre del organismo tributario que aquí se utiliza es ficticio.

FIGURA 9.8
(Continuación)

3. La Dirección General de Impuestos Nacionales es muy lenta en la devolución de impuestos pagados en exceso.	1. Muy de acuerdo	2. De acuerdo	3. Ni de acuerdo, ni en desacuerdo	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	4. De acuerdo	3. Ni de acuerdo, ni en desacuerdo	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	4. De acuerdo	3. Ni de acuerdo, ni en desacuerdo	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	2. De acuerdo	3. Ni de acuerdo, ni en desacuerdo	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 denomina *escala aditiva*. La figura 9.9 muestra un ejemplo de cómo calificar una escala de Likert.

Una puntuación se considera alta 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:

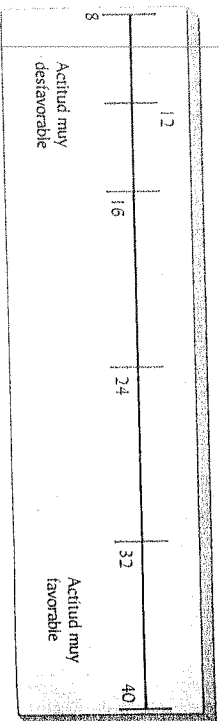
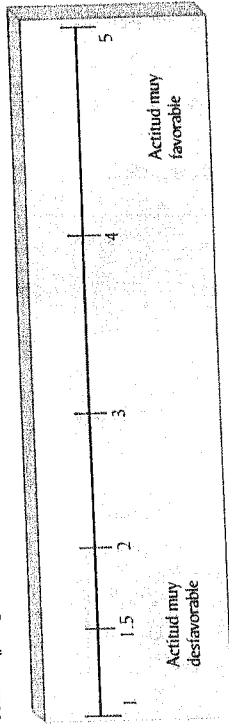


FIGURA 9.9
Ejemplo de cómo calificar una escala Likert.

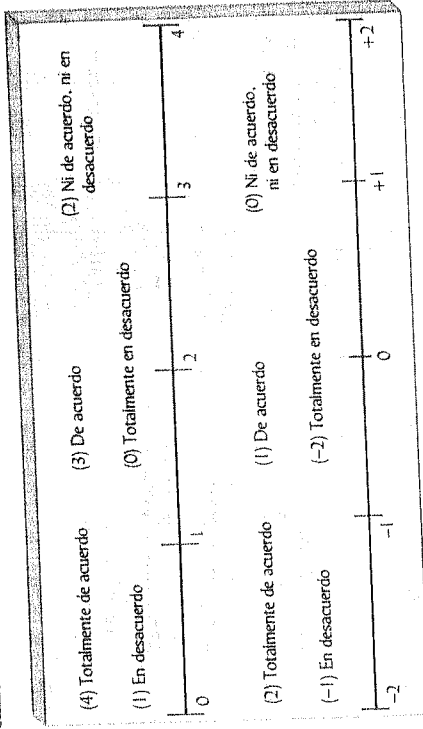
1. El personal de la Dirección General de Impuestos Nacionales es gansoso al atender al público.	<input checked="" type="checkbox"/> Muy de acuerdo	<input type="checkbox"/> De acuerdo	<input type="checkbox"/> Ni de acuerdo, ni en desacuerdo	<input type="checkbox"/> En desacuerdo	<input type="checkbox"/> Muy en desacuerdo
2. La Dirección General de Impuestos Nacionales se caracteriza por la deshonestedad de sus funcionarios.	<input checked="" type="checkbox"/> Muy de acuerdo	<input type="checkbox"/> De acuerdo	<input type="checkbox"/> Ni de acuerdo, ni en desacuerdo	<input type="checkbox"/> En desacuerdo	<input type="checkbox"/> Muy en desacuerdo
3. Los servicios que presta la Dirección General de Impuestos Nacionales son en general muy buenos.	<input type="checkbox"/> Muy de acuerdo	<input checked="" type="checkbox"/> De acuerdo	<input type="checkbox"/> Ni de acuerdo, ni en desacuerdo	<input type="checkbox"/> En desacuerdo	<input type="checkbox"/> Muy en desacuerdo
4. La Dirección General de Impuestos Nacionales informa claramente sobre cómo, dónde y cuándo pagar los impuestos.	<input type="checkbox"/> Muy de acuerdo	<input type="checkbox"/> De acuerdo	<input checked="" type="checkbox"/> Ni de acuerdo, ni en desacuerdo	<input type="checkbox"/> En desacuerdo	<input type="checkbox"/> Muy en desacuerdo
5. La Dirección General de Impuestos Nacionales es muy lenta en la devolución de impuestos pagados en exceso.	<input checked="" type="checkbox"/> Muy de acuerdo	<input type="checkbox"/> De acuerdo	<input type="checkbox"/> Ni de acuerdo, ni en desacuerdo	<input type="checkbox"/> En desacuerdo	<input type="checkbox"/> Muy en desacuerdo
6. La Dirección General de Impuestos Nacionales informa oportunamente sobre cómo, dónde y cuándo pagar los impuestos.	<input type="checkbox"/> Muy de acuerdo	<input checked="" type="checkbox"/> De acuerdo	<input type="checkbox"/> Ni de acuerdo, ni en desacuerdo	<input type="checkbox"/> En desacuerdo	<input type="checkbox"/> Muy en desacuerdo
7. La Dirección General de Impuestos Nacionales tiene normas y procedimientos bien definidos para el pago de impuestos.	<input type="checkbox"/> Muy de acuerdo	<input type="checkbox"/> De acuerdo	<input checked="" type="checkbox"/> Ni de acuerdo, ni en desacuerdo	<input type="checkbox"/> En desacuerdo	<input type="checkbox"/> Muy en desacuerdo
8. La Dirección General de Impuestos Nacionales tiene malas relaciones con la gente porque cobra impuestos muy altos.	<input checked="" type="checkbox"/> Muy de acuerdo	<input type="checkbox"/> De acuerdo	<input type="checkbox"/> Ni de acuerdo, ni en desacuerdo	<input type="checkbox"/> En desacuerdo	<input type="checkbox"/> Muy en desacuerdo

Valor = 1 + 2 + 1 + 3 + 1 + 1 + 2 + 1 = 12

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 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 ($12/8 = 1.5$):



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 de -2 a +2 en lugar de 1 a 5. Pero esto no importa porque se cambia el marco de referencia de la interpretación. Veamoslo gráficamente.



Simplemente 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 se 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	(1) En desacuerdo
(3) De acuerdo	(2) Ni de acuerdo, ni en desacuerdo	(5) Indeciso, pero más bien de acuerdo
(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 encuestados 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 correlacionan 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 (Pantlagua, 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...?)

	Indispen- sable (5)	Sumamente importante (4)	Mediana- mente importante (3)	Poco importante (2)	No se toma en cuenta (1)
— Precio					
— Forma de pago (contado-credito)					
— Tiempo de entrega					
— Lugar de entrega					
— Garantía del producto					
— Servicio de reparación					
— Prestigio del producto (marca)					
— Prestigio de la empresa proveedora					
— Comunicación que se tienen con la(s) persona(s) que representan al proveedor					
— Apoyo del proveedor a los requerimientos legales del producto					
— Cumplimiento del proveedor con las especificaciones					

FIGURA 9.10 (Continuación).

	Indispen- sable (5)	Sumamente importante (4)	Mediana- mente importante (3)	Poco importante (2)	No se toma en cuenta (1)
— 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.

Maneras de aplicar la escala Likert

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 necesario que se le entregue al entrevistado 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:

	Indispen- sable	Sumamente importante	Medianamen- te importante	Poco importante	No se toma en cuenta
— 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					

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 evaluará cuidadosamente en la prueba piloto.

Diferencial semántico¹²

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 debe 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"
 justo: : injusto

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.

Osgood, Suci y Tannenbaum (1957) nos indican que, si el respondiente considera que el objeto de actitud se relaciona muy estrechamente con uno u otro extremo de la escala, la respuesta se marca así:

justo: X : injusto
 o de la siguiente manera: X : injusto

Si el respondiente considera que el objeto de actitud se relaciona estrechamente con uno u otro extremo de la escala, la respuesta se marca así (dependiendo del extremo en cuestión):

¹² Para profundizar en el diferencial semántico se recomienda consultar Osgood, Suci y Tannenbaum (1957, 1976a) y (1976b); así como Heise (1976).

justo: X : injusto
 justo: X : injusto

Si el respondiente considera que el objeto de actitud se relaciona mediante con alguno de los extremos, la respuesta se marca así (dependiendo del extremo en cuestión):

justo: X : injusto
 justo: X : injusto

Y si el respondiente considera que el objeto de actitud ocupa una posición neutral en la escala (ni justo ni injusto en este caso), la respuesta se marca así:

justo: X : injusto

Es decir, en el ejemplo, cuanto más justo considere al candidato "A" más me acerco al extremo "justo", y viceversa, entre más injusto lo considero más me acerco al extremo opuesto.

Algunos ejemplos de adjetivos se muestran en la figura 9.11.

La figura 9.11 presenta solo algunos ejemplos, desde luego hay muchos más que han sido utilizados o que pudieran pensarse. La elección de adjetivos depende del objeto de actitud a calificar, los adjetivos deben poder aplicarse a éste.

Codificación de las escalas

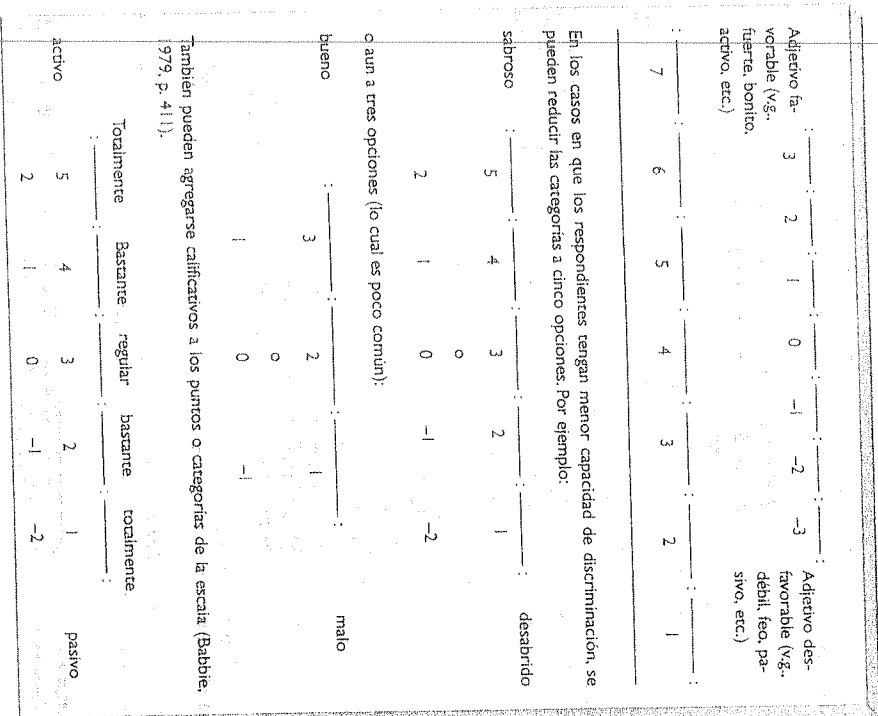
Los puntos o categorías de la escala pueden codificarse de diversos modos, que se presentan en la figura 9.12.

FIGURA 9.11

Ejemplos de adjetivos bipolares.

fuerte-débil	potencioso-impotente
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	arriba-abajo
bueno-malo	útil-ínutil
dulce-ácido	favorable-desfavorable
profundo-superficial	agresivo-tímido

FIGURA 9.12
Maneras comunes de codificar el diferencial semántico.

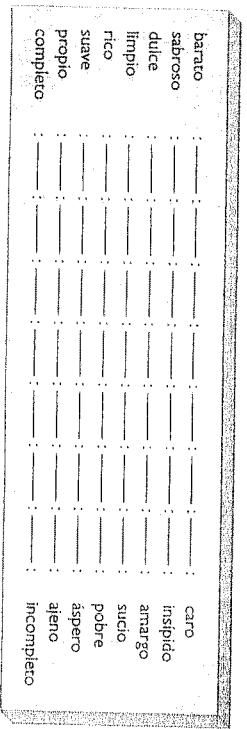


Codificar de 1 a 7 o de -3 a 3 no tiene importancia, siempre y cuando estos sean consistentes del marco de interpretación. Por ejemplo, si una persona calificó al objeto de actitud candidato "A" en la escala justo-injusto, marcando la categoría más cercana al extremo "injusto", la calificación puede ser "1" o "-3".

Justo: 7 6 5 4 3 2 1 X injusto

Justo: 3 2 1 0 -1 -2 -3 X injusto

FIGURA 9.13
Ejemplo parcial de un diferencial semántico para medir la actitud hacia un producto comestible.



Pero en un caso la escala oscila entre 1 y 7 y en el otro caso entre -3 y 3. Si deseamos evitar el manejo de números negativos utilizamos la escala de 1 a 7.

Maneras de aplicar el diferencial semántico

La aplicación del diferencial semántico puede ser *autoadministrada* (se le proporciona la escala al sujeto y este marca la categoría que describe su reacción o considera conveniente) o *mediante entrevista* (el entrevistador marca la categoría que corresponde a la respuesta del sujeto). En esta segunda situación es muy conveniente mostrar una tarjeta al respondiente, que incluya los adjetivos bipolares y sus categorías respectivas.

La figura 9.13 muestra parte de un ejemplo de un diferencial semántico utilizado en una investigación para evaluar la actitud hacia un producto (Comunicometría, 1988).

Las respuestas se califican de acuerdo con la codificación. Por ejemplo, si una persona tuvo la siguiente respuesta:

rico: X : pobre

Y la escala oscila entre 1 y 7, esta persona obtendría un siete (7).

En ocasiones se puede incluir la codificación en la versión que se les presenta a los respondientes con el propósito de clarificar las diferencias entre las categorías.

Por ejemplo:

sabroso: : insípido

Pasos para integrar la versión final

Para integrar la versión final de la escala se deben llevar a cabo los siguientes pasos:

1. *Generamos una lista de adjetivos bipolares exhaustiva y aplicable al objeto de actitud a medir.* De ser posible, resulta conveniente que se seleccionen adjetivos que hayan sido utilizados en investigaciones similares a la nuestra (con textos parecidos).

1. **Construimos una versión preliminar** de la escala y la administramos a un grupo de sujetos a manera de prueba piloto.
2. **Correlacionamos las respuestas de los sujetos para cada par de adjetivos o ítem**. Así, correlacionamos un ítem con todos los demás (cada par de adjetivos contra el resto).
3. **Calculamos la confiabilidad y validez de la escala total** (todos los pares de adjetivos).
4. **Seleccionamos los ítems que presenten correlaciones significativas** con los demás ítems. Naturalmente, si hay confiabilidad y validez, estas correlaciones serán significativas.
5. **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.

Su interpretación depende del número de ítems o pares de adjetivos. Asimismo, en ocasiones se califica el promedio obtenido en la escala total.

Puntuación total (Número de ítems). 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 Guttman³²

Este método para medir actitudes fue desarrollado por Luis Guttman. Se basa en el principio de que algunos ítems indican en mayor medida la fuerza o intensidad

sabroso	:	X	:	:	:	:	:	insípido
rico	:	X	:	:	:	:	:	pobre
suave	:	:	X	:	:	:	:	áspero
balanceado	:	:	:	X	:	:	:	desbalanceado

Valor = 6 + 7 + 6 + 6 = 25

FIGURA 9.14 Ejemplo de cómo calificar un diferencial semántico.

dad 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 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).

Para construir el escalograma es necesario desarrollar un conjunto de afirmaciones pertinentes al objeto de actitud. Estas deben variar de 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", "sí-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:

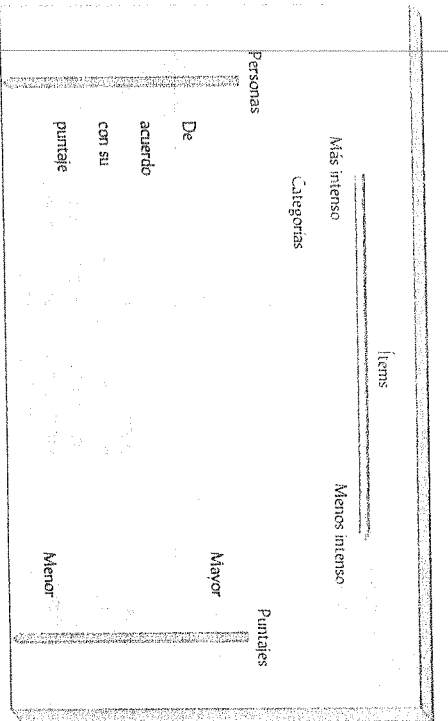
1. **Obtener el puntaje total** de cada sujeto en la escala.
2. **Ordenar a los sujetos de acuerdo con su puntaje total** (del puntaje mayor al menor, de manera vertical descendente).
3. **Ordenar a las afirmaciones de acuerdo con su intensidad** (de mayor a menor y de izquierda a derecha).
4. **Construir una tabla donde se crucen los puntajes de los sujetos ordenados con los ítems y sus categorías jerarquizados**. 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.

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.

³² Para profundizar en esta escala se sugiere consultar Nie, et al. (1975), Black y Champion (1976), Guttman (1976), Lingoes (1976), Dotson y Sumers (1976) y Pictua (1979).

FIGURA 9.15
Manera de disponer los resultados para el análisis de ítems mediante la técnica de Cornell.



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 valor de izquierda a derecha. Hay que recordar que si la afirmación es

SUJETO	AFIRMACIONES								PUNTAJES TOTALES
	A		B		C		D		
	DA	ED	DA	ED	DA	ED	DA	ED	
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	1
12		X		X		X		X	1
13		X		X		X		X	0
14		X		X		X		X	0

DA = De acuerdo o I, ED = En desacuerdo o 0

negativa, la codificación se invierte ("La calidad es poco importante para el desarrollo de una empresa", "de acuerdo" se codifican 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, muy probablemente también lo estén 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 están "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 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 la afirmación más intensa, también lo están 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 avión 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 según su 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 del 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 la escala. 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 intensas). 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.

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.

No se aprecia ninguna inconsistencia. En cambio, en la tabla 9.4 se aprecian cuatro inconsistencias o errores, las respuestas están desubicadas respecto a los puntos de corte, rompen el patrón de intensidad (los errores están encerrados en un círculo).

Como se ha 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:

$$\text{Coeficiente de reproductividad} = \frac{\text{Número de errores o inconsistencias}}{\text{Número total de respuestas}}$$

Tabla 9.2
Ejemplo de errores o inconsistencias en un escalograma de Guttman.

SUJETO	AFIRMACIONES												PUNTAJES 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			1
6										X			0

Tabla 9.3
Ejemplo de establecimiento de los puntos de corte en la técnica de Cornell.

SUJETO	AFIRMACIONES												PUNTAJES TOTALES
	A			B			C			D			
	DA	ED		DA	ED		DA	ED		DA	ED		
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			1
12		X		X			X			X			1
13		X		X			X			X			0
14		X		X			X			X			0

DA = De acuerdo o I. ED = En desacuerdo o 0

Tabla 9.4
Ejemplo de errores respecto a los puntos de corte.

SUJETO	AFIRMACIONES												PUNTAJES TOTALES
	A			B			C			D			
	DA	ED		DA	ED		DA	ED		DA	ED		
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			2
8		X		X			X			X			2
9		X		X			X			X			2
10	(X)									X			2
11	(X)									X			1
12	(X)									X			1
13		X		X			X			X			0
14		X		X			X			X			0

--- = puntos de corte (líneas puntuadas)

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 es menor que .90 no se acepta la escala. Originalmente Guttman 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 a 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

Cuando 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, todos los ítems deben tener el mismo número de categorías de respuesta. Éste 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 escalograma de Guttman es una escala estrictamente ordinal pero que se suele usar como si fuera de intervalo. Puede aplicarse mediante entrevista (con uso de tarjetas que contengan las opciones o categorías de respuesta) o puede ser autoadministrada.

9.5.2 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 mida. 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 ellas. 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, específica

De la hemeroteca:

- No entré
- A consultar algún periódico
- A usar las videocaseteras
- A estudiar

- A hacer trabajos
- A sacar copias
- A leer algún libro
- Otros, específica

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 localización de material
- Otros, específica

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?*

- LEM
- BCC
- MME

O bien debe asignar un puntaje a una o diversas cuestiones.

Ejemplo

¿CUANTO 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.
- _____ 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).

* Conductores ficticios.

_____	Sistemas de información y control.
_____	Auditoría administrativa.
_____	Planificación estratégica.
_____	Sistemas de computación.
_____	Mercadeotecnia y comercialización.
_____	Otros (especificar):

En otras se anota una cifra: _____

¿Cuántas casas habitación terminadas existen en el municipio de Aizcorbe?

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:

<i>Ejemplo</i>	
_____	¿Por qué asiste a psicoterapia?

_____	¿Qué opina del programa de televisión <i>Los cazadores</i> ?

_____	¿De que manera la directiva de la empresa ha logrado la cooperación del sindicato para el proyecto de calidad?

¿Conviene usar 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 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. Éstos 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 completo. 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 las personas tienen en mente, no siempre se captura lo que pasa por la cabeza 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 debe 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 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 (planeamiento 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: ¿cuál aproximadamente es su nivel mensual de ingresos? y preguntando: ¿cuántos focos eléctricos tiene aproximadamente en su

casas. 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, 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:

- 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.
1. Sólo a quienes distribuyen selectivamente datos sobre calidad. ¿A qué niveles de la empresa?

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, presencia de focos en el jardín de la casa, candiles y otros factores.

Sólo a quienes distribuyen selectivamente datos sobre calidad. ¿A qué funciones?

¿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 en los receptores o usuarios de los datos en aspectos del control de calidad distribuidos selectivamente. Se justifica hacer estas dos preguntas, pues 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 son 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. 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 (como lo hacemos con las escalas de actitudes).

Ejemplos de preguntas precodificadas

¿Tiene usted inversiones en la Bolsa de Valores?

- Si
- No

Cuando 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 menos tiempo.

¿Que características debe tener una pregunta?

Independientemente de que las preguntas sean abiertas o cerradas y de que 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?, tiende 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., ¿cual es su tipo de bebida favorita?, etcétera). Y hay temáticas 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 aun otras formas de medición. Tal es el caso de temas como homosexualismo, prostitución, pornografía, anticonceptivos y la drogadicción.
- C. *Las preguntas deben referirse preferentemente a un solo aspecto o relación lógica.* Por ejemplo, la pregunta: ¿acostumbra usted ver televisión y escuchar radio diariamente?, expresa dos aspectos y 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 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?
- E. *Las preguntas no pueden apoyarse en instituciones, ideas respaldadas socialmente ni en evidencia comprobada.* Es también una manera de inducir res-

Súmmamente productivos	Más bien productivos	Más bien improductivos	Súmmamente improductivos
------------------------	----------------------	------------------------	--------------------------

puestas. Por ejemplo, la pregunta: La Organización Mundial de la Salud ha realizado diversos estudios y condujo 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 solo tiene que elegir una, puede ocurrir que el orden en que se presentan dichas alternativas afecte las respuestas de los sujetos* (v.g., tienden 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 lograra disminuir verdaderamente la inflación? y el 33.33% de la 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 apropiado 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 adelantándose en la situación. No se recomienda comenzar con preguntas difíciles o muy directas. Imaginemos un cuestionario diseñado para obtener opiniones en torno al aborto que empiece con una pregunta poco sutil tal como: ¿Está 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 e institución. Pero en otras ocasiones es mucho 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. Estas 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).

¿Tiene este ejido o comunidad, ganado, aves o colmenas que sean de propiedad colectiva? (CIRCULE LA RESPUESTA)

No 2

(pase a 30)

Si 1

(continúe)

¿Se ha obtenido la cooperación de todo el personal o la mayoría de éste para el proyecto de calidad?

2 No

(pase a la pregunta 27)

1 Si

(pase a la pregunta 26)

	Comp. (5)	Bast. (4)	Reg. (3)	Poco (2)	Nada (1)
1. ¿Las políticas de su proveedor?					
2. ¿Sus finanzas (estado financiero)?					
3. ¿Los objetivos de su área de ventas?					
4. ¿Sus programas de capacitación para vendedores?					
5. ¿Número de empleados de su área de ventas?					
6. ¿Problemas laborales?					
7. ¿Los métodos de producción que tienen?					
8. ¿Otros clientes de ellos?					
9. ¿Su índice de rotación personal?					

FIGURA 9.14
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 e incluidas en la tesis profesional, pero nunca se comunicarán 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)

COMUNICACIÓN ESTÁ HACIENDO UNA ENCUESTA CON EL PROPÓSITO DE CONOCER LAS OPINIONES QUE SE TIENEN ACERCA DE ESTA EMPRESA. Y PARA ELLO LE PEDIRIA FUERA TAN AMABLE DE CONTESTAR UNAS PREGUNTAS. NO LE TOMARÁ MÁS DE 30 MINUTOS. LA INFORMACIÓN QUE NOS PROPORCIONE SERÁ MANEJADA CON LA MÁXIMA ESTRICTA CONFIDENCIALIDAD. DESDE LUEGO, NO HAY PREGUNTAS DELICADAS.

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 *cartilla 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. En la figura 9.16 se presentan algunos textos ilustrativos de cartas introductorias a un cuestionario.

La manera en que pueden distribuirse las preguntas, categorías de respuestas e instrucciones es variada. Algunos prefieren colocar las preguntas a la iz-

quiere y las respuestas a la derecha, con lo que se tendría un formato como el siguiente:

¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Otros dividen el cuestionario por secciones de preguntas y utilizan un formato horizontal.

Presentación				
Preguntas sobre motivación intrínseca:				
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Preguntas sobre satisfacción laboral				
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
¿ Como ... ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Etc.				

Otros 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; que las instrucciones, preguntas y respuestas se diferencien que el formato no resulte visualmente tedioso y se pueda leer sin dificultad.

¿De qué tamaño debe ser un cuestionario?

No existe una regla al respecto, pero, como menciona Padua (1979), si es muy corto se pierde información, y si resulta largo puede resultar tedioso. En este último caso, las personas pueden negarse a responder o, al menos, lo responderán en forma incompleta. 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 apartado). Cuestionarios que duren más de 30 minutos deben resultar tediosos a menos que los respondientes estén muy motivados para contestar (V.G.). cuestionarios de personalidad, cuestionarios para obtener un trabajo, *line recommendation* que puede ayudarnos a evitar un cuestionario más largo de lo requerido son: "No hacer 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 darles 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 sigue el siguiente procedimiento, basado parcialmente en Rojas (1981, pp. 150-151):

1. Seleccionar determinado número de cuestionarios mediante un muestreo adecuado de nuestro, 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 alejadas en remas, 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 Comunicmería (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 respuesta con mayor frecuencia de mención)	Número de frecuencia de medición
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, estos 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: ¿Que sugerencias podría hacer para mejorar al programa *Exceler*? 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".

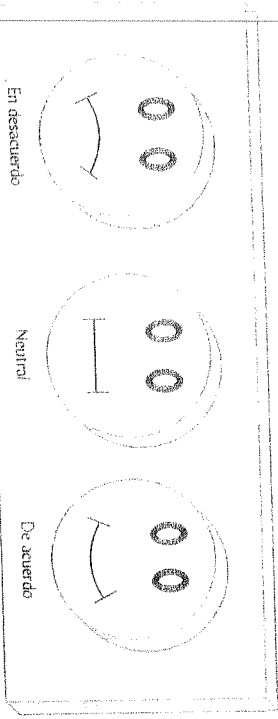
¿En qué contextos puede administrarse o aplicarse un cuestionario?

Los cuestionarios pueden ser aplicados de diversas maneras:

1. **Autoadministrado.** El cuestionario se 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.
2. **Por entrevista personal.** 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.
3. **Por entrevista telefónica.** Esta situación es similar a la anterior, solo 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.
4. **Autoadministrado y enviado por correo postal, electrónico o servicio de mensajería.** Los respondientes contestan directamente el cuestionario, ellos marcan o anotan las respuestas, no hay intermediario. Pero los cuestionarios no se entregan 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 efectos que difícilmente puedan dedicarse a un solo sujeto 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 muchas cuantías preguntas (su administración toma entre 4 y 5 minutos), 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 las preguntas. Cuando un cuestionario o escala es aplicada de forma masiva suele denominarse "encuesta".

Cuando se utiliza la entrevista telefónica se debe tomar en cuenta el horario. Ya que si hablamos solo a una hora (digamos en la mañana), nos encontraremos con unos cuantos subgrupos de la población (Vd., un caso de caso).

Cuando lo enviamos por correo o es autoadministrado directamente, las instrucciones deben ser muy precisas, claras y completas. Y debemos dar instrucciones que motiven al respondiente para que continúe contestando el cuestionario (Vg., 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 Giffel (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 agrada-

bles suelen resultar excelentes entrevistadoras. 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 mostró sincero, la manera como respondió, el tiempo que duró la entrevista, el lugar donde se realizó, las características del entrevistado, los contrastes que se presentaron y la manera en que 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 dispone, el tiempo de entrega de los resultados, los objetivos de la investigación y el tipo de respondientes (edad, nivel educativo, etcétera).

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).

Cuando los cuestionarios son muy complejos de contestar o aplicar, suele utilizarse un manual que explica a fondo las instrucciones y como debe responderse o ser administrado.

¿Cuál es el proceso para construir un cuestionario?

Siguiendo los pasos para construir un instrumento de medición rendiríamos 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, todos deberán ajustarse a esto. Si no se es consistente, algunos respondientes o entrevistadores pueden desconcertarse.

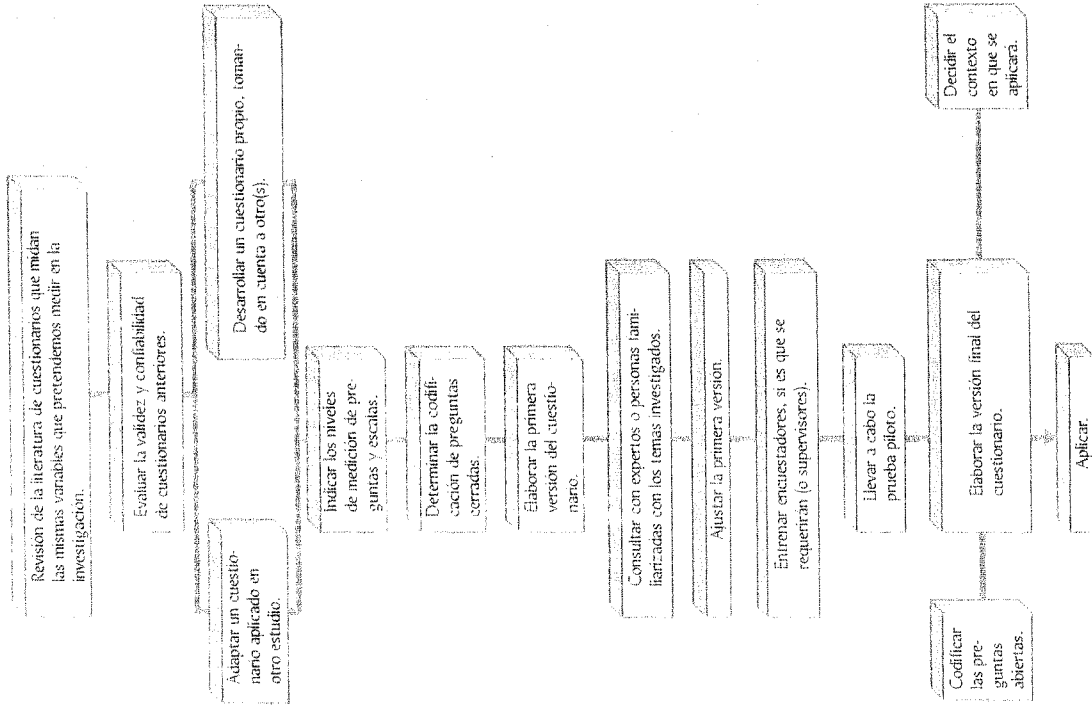
9.6.3 Análisis del contenido

¿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 cómo 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 radifónicos, artículos en prensa, libros, poemas, conversaciones, pinturas, discursos, cartas, melodías, reglamentos, etcétera). Por ejemplo, puede

Proceso para construir un cuestionario.



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; penetrarse 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. Revelar diferencias en el contenido de la comunicación (entre personas, grupos, instituciones, países).
3. Comparar mensajes, niveles y medios de comunicación.
4. Auditar el contenido de la comunicación y compararlo contra estándares u objetivos.
5. Construir y aplicar estándares de comunicación (políticas, normas, etcétera).
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 entre otras cosas, 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; 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 calificación, el proceso en virtud 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 calificar es necesario definir el universo, las unidades de análisis y las categorías de análisis.

El universo

El universo podría ser la obra completa de Franz Kafka, las emisiones de un noticiario 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 el ciclo escolar, los discursos pronunciados por varios congresistas 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

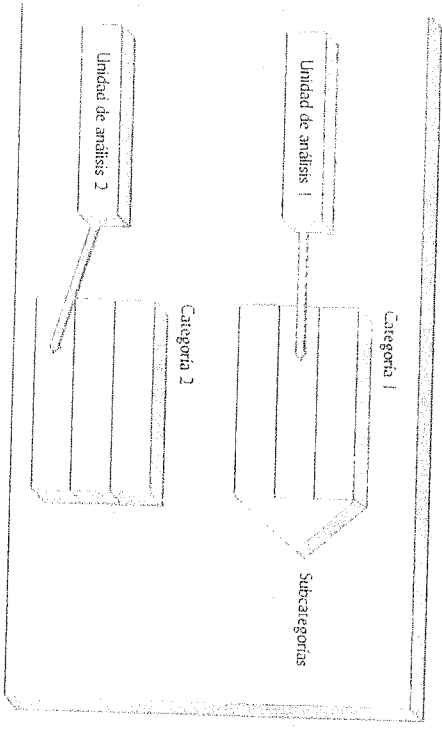
Las unidades de análisis constituyen segmentos del contenido de los mensajes que son caracterizados para ubicarlos dentro de las categorías. Berelson (1952) menciona cinco unidades importantes de análisis:

- La palabra.** Es la unidad de análisis más simple, aunque como señala Kerlinger (1975), puede haber unidades más pequeñas como letras, fonemas, símbolos. Así, se puede medir cuantas veces aparece una palabra en un mensaje (v.g., veces que en un programa televisivo de fin de año se menciona al presidente).
- El tema.** 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 autorreferencia. Este sería el tema más extenso. Los temas que constituyen este podrían definirse como cualesquiera oraciones que usen "yo", "mi" y otros términos que indiquen referencia al yo del escritor. Así, se analizaría que tanta autorreferencia contienen las cartas.

Si los temas son complejos, el análisis del contenido es más difícil, especialmente si se complica al incluirse más de una oración simple. El tema. 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 items 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 ese caso lo que se analiza es el material simbólico total. El personaje. Un individuo, un personaje televisivo, un líder histórico, etcétera. Aquí lo que se analiza es el personaje. Medidas de espacio-tiempo. Son unidades físicas como el centímetro-columna (por ejemplo, en la prensa), la línea (en escrito), el minuto (en una conversación o en radio), el período de 20 minutos (en una interacción), el cuadro (en televisión), cada vez que se haga una pausa (en un discurso).

Las unidades se insertan, colocan o caracterizan en categorías, esto podría representarse de la siguiente manera:



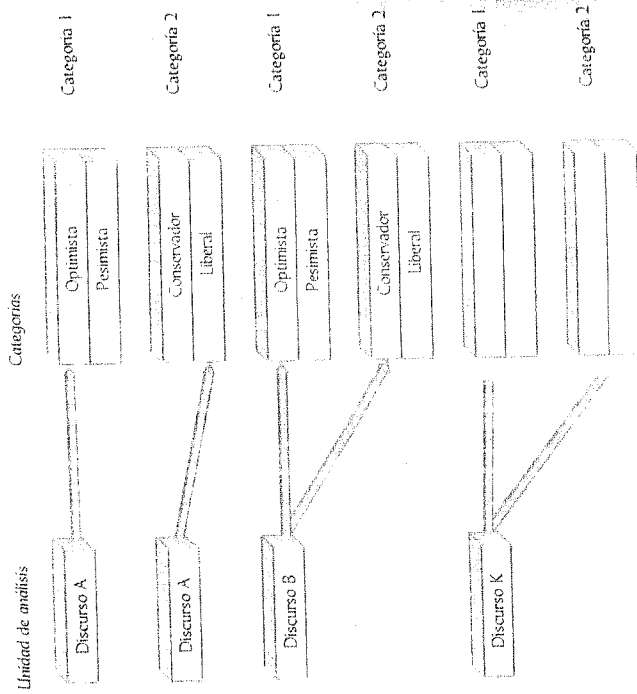
Respecto a la pregunta: ¿qué unidad debe seleccionarse? Esto depende de los objetivos y preguntas de investigación. Sin embargo, Berelson (1952) sugiere lo siguiente:

- En un solo estudio se puede utilizar más de una unidad de análisis.
- Los cálculos de palabras y las unidades amplias, como el item y las medidas de espacio-tiempo, son más adecuadas en los análisis que dan énfasis a asuntos definidos.
- Las unidades amplias y las más delimitadas son válidas para la aceptación o rechazo en una categoría.
- Las unidades amplias generalmente requieren de menos tiempo para su codificación que las unidades pequeñas, referidas a las mismas categorías y materiales.

- 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.
- El tema es adecuado para análisis de significados y las relaciones entre éstos.

Categorías

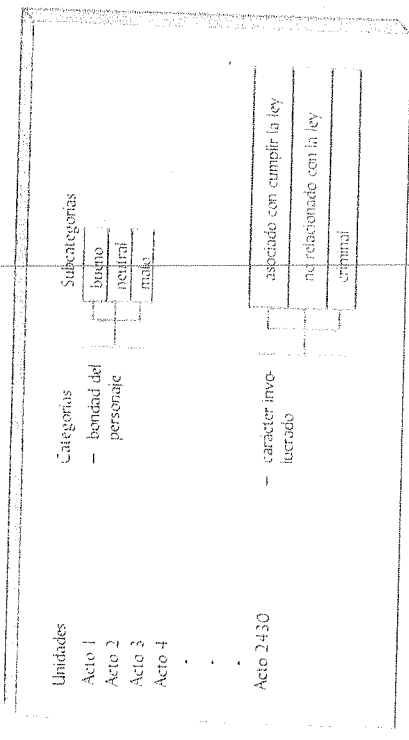
Las categorías son los niveles donde serán caracterizadas las unidades de análisis. Como menciona Holsti (1968), son las "cajitas o cajones" en las cuales van clasificadas las unidades de análisis. Por ejemplo, un discurso podría clasificarse como optimista o pesimista, como liberal o conservador. Un 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 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:



La selección de categorías también depende del planteamiento del problema.

Tipos de categorías

Krippendorff (1982) señala cinco tipos de categorías:

- De asunto o tópico. Se refieren a cual es el asunto, tópico o tema tratado en el contenido (¿de qué trata el mensaje o la comunicación?).

Análizar el último informe del secretario o ministro de Hacienda o Finanzas.

- Categorías: Tema financiero
- Subcategorías: Deuda
- Impuestos:
- Planeación haciendaria
- Inflación
- Electricidad

De atención: Estas categorías se refieren a cómo es tratado el asunto (¿positiva o negativamente?, favorable o desfavorable?, nacionalista o no nacionalista, etc.).
Ejemplo: Comparar la manera en que dos noticieros televisivos hablan de la posibilidad de una moratoria unilateral en el pago de la deuda externa de Latinoamérica.

- Categorías: Tono en el tratamiento de la deuda externa
- Subcategorías: A favor de la moratoria unilateral.
- En contra
- Neutral

De valores: Se refieren a categorías que indican qué valores, intereses, metas, deseos o creencias son revelados.
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).

- Categorías: Ideología del esposo.
- Subcategorías: Muy tradicional
- Más bien tradicional
- Neutral
- Más bien liberal
- Muy liberal

De receptores: Estas categorías se relacionan con el destinatario de la comunicación (¿a quien van dirigidos los mensajes?).

Receptores:
Analizar a quienes se dirige más un líder sindical en sus declaraciones a los medios de comunicación durante un periodo determinado.

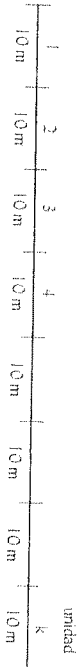
Categorías: Receptores a quienes dirige el mensaje.
Subcategorías: Opinión pública en general

- Presidente
- Cabine económico
- Cabine agropecuario
- Gobierno en general
- Sector empresarial
- Obreros afiliados a su sindicato
- Obreros no afiliados a su sindicato
- Obreros en general (afiliados y no afiliados)
- Electricidad

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 con una película, etc. No se deben 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 se van a considerar la unidad de análisis.



dirección de la telenovela

- 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, que deben cumplir 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.

Pero las categorías no siempre son mutuamente excluyentes. Por ejemplo, al analizar comerciales televisivos podríamos tener, entre otras categorías, las siguientes:

Categorías	Subcategorías
Presencia de color	color — blanco-negro — mixto
Necesidad fisiológica a que se apela principalmente	Hambre — Sed — Sexo — Respiración (oxígeno) — Abrigo — Sueño

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".

En ciertos casos especiales, puede interesar al analista un sistema de categorías donde éstas no sean mutuamente excluyentes. Pero no es lo normal.

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 se acudirá a un estudio de Greenberg, Edison, Korzeny, Fernandez-Collado y Atkin (1980). El estudio consistió en un análisis de contenido de las series televisadas por las tres grandes cadenas de Estados Unidos: ABC, CBS y NBC. Se analizaron diversos programas durante tres períodos de 1975 a 1978, para evaluar el grado en que la televisión estadounidense mostraba actos prosociales y antisociales como medida de la violencia televisada. Las categorías y subcategorías eran las siguientes:

El estudio consideró como unidad de análisis 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 que se va a comprender en cada caso y que 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 categorías. Se cuenta cuántas veces se repite cada categoría o subcategoría (cuantas unidades de análisis entraron en cada una de las categorías). Por ejemplo, Greenberg et al. (1980, p. 113) encontraron los resultados que se muestran en la Tabla 7.5

TABLA 7.5
Ejemplo de los resultados de la codificación de acuerdo con el estudio de Greenberg et al. (1980).

INCIDENCIA DE ACTOS ANTISOCIALES EN LOS TRES PERIODOS	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. Amenaza 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		2 713	
Horas analizadas	(68.5)		(58)		(63)	

* Año 1 incluye de octubre de 1975 hasta que se grabó un episodio de (1976), año 2 igual pero en 1976-1977 y año 3 igual pero de 1977-1978.
 f = número de casos o unidades.
 % = porcentajes.

Ejemplo de categorías:

Actos prosociales	<p>Atruismo (definido como un acto en el cual humanos o animales comparten, se ayudan y cooperan, siempre y cuando no participen en un acto ilícito).</p> <p>Mostrar afecto (referring to offering or exhibiting overtly emotions positive hacia humanos y animales).</p> <p>Explicación de los sentimientos de uno mismo (consiste en afirmaciones verbales que explican los sentimientos, pensamientos o acciones de uno mismo ("YO"), y que tienen como propósito lograr un resultado positivo).</p> <p>Explicación de los sentimientos de uno y otros (consiste en afirmaciones verbales que explican los sentimientos, pensamientos o acciones de los demás y que tienen como propósito lograr un resultado positivo).</p>
Actos antisociales	<p>Agresión física (preferida a cualquier conducta manifiesta cuya intención es dañar, asustar, lesionar o injuriar a uno mismo o a otro individuo, animal o propiedad).</p> <p>Asalto con algún objeto (arma, un libro, un lápiz, etc.)</p> <p>Asalto sin un objeto (v.g., utilizando las manos).</p> <p>Amenaza física (ponen en peligro el bienestar o la integridad de alguien).</p> <p>Disparar a alguien con un arma de fuego.</p>
Actos antisociales	<p>Agresión verbal (el envío de mensajes nocivos tales como rechazos, amenazas verbales, insultos o frases hostiles).</p> <p>Robo (la acción de tomar intencional y deliberadamente una propiedad que pertenece a otra persona o institución, sin derecho o permiso). En-gaño (engaño intencional de alguien con el propósito de provocar en detrimento de un individuo, grupo o institución (v.g., fraude, mentiras, asumir la identidad de otros).</p>

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 presenten a las variables de la investigación.

Los demás pasos son:

- 4. **Seleccionar las codificadoras:** Los codificadores son las personas que habrán de asignar las unidades de análisis a las categorías. Deben tener 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 (años, número o equivalente), ubicación, formato (en centímetros).

En la columna, nombre del codificador, día de codificación, etcétera. Una hoja puede servir para una o varias unidades, según 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". El 1.0 es "atruismo", 1.1 "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 sus categorías.

5. **Proporcionar entrenamiento de codificadoras.** Este entrenamiento incluye que los codificadoras se familiaricen y comprendan 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.

FIGURA 2.14

Elementos de una hoja de codificación.

Indicador de quién codificó número, iniciales o letras		Frecuencias (veces que se repite cada categoría)		Descripción del material que se analizará (discurso, nota perio- dística, sesión terapéutica, etc.)	
FECHA: 29-XI-88	CODIFICADOR: I	MATERIAL A ANALIZAR: DURACIÓN: 30 MINUTOS	FRECUENCIAS	CAPÍTULO DE LA SERIE "A1" (CBS)	TOTALES
Fecha en que se codificó el material	CODIFICADOR: I	MATERIAL A ANALIZAR: DURACIÓN: 30 MINUTOS	FRECUENCIAS	CAPÍTULO DE LA SERIE "A1" (CBS)	TOTALES
CATEGORÍAS					
A C T O S					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
A GRESIÓN 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
A GRESIÓN 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

Total de frecuencias o unidades

Al reverso pueden solicitarse comentarios del codificador.

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 llamado "confiabilidad intracodificador". La cual mide la estabilidad de la prueba y reprobada 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 logren correctamente 20, la confiabilidad será de $\left(\frac{20}{20}\right)$ 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 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 considerará un caso poco confiable.

Ejemplo

Codificador A	0.89	Codificador B	0.92	Codificador C	0.67	Codificador D	0.67
---------------	------	---------------	------	---------------	------	---------------	------

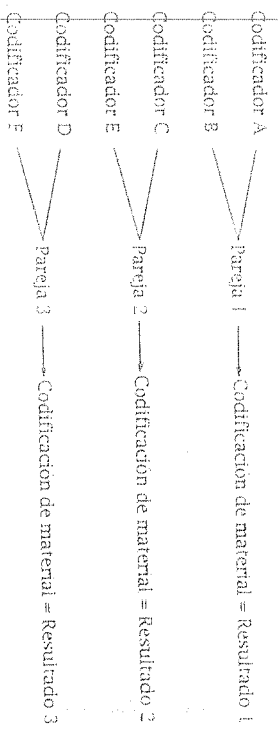
El codificador "D" tiene baja confiabilidad

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:

$$\text{Confiabilidad} = \frac{\text{Número total de acuerdos entre dos parejas}}{\text{Número total de unidades de análisis codificadas entre parejas}}$$

Después se suman los resultados de esta fórmula y se divide entre el número de comparaciones, que depende del número de parejas. Veamoslo 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

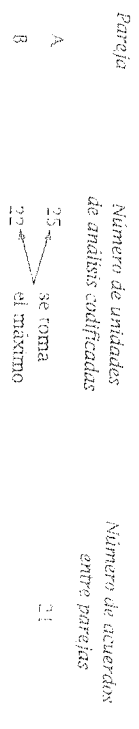
Se aplica la fórmula de confiabilidad entre parejas

$$C_{1 \times 2} = \frac{17}{18} = 0.94$$

$$C_{1 \times 3} = \frac{16}{18} = 0.88$$

$$C_{2 \times 3} = \frac{16}{18} = 0.88$$

Debe observarse que no hubo consenso total entre cuantas 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 de unidades de análisis codificadas por alguna de las parejas. Si fueran:



La fórmula sería $C_{AB} = \frac{21}{25} = 0.84$

Se obtiene la confiabilidad total (que es la suma de las confiabilidades entre parejas sobre el número de comparaciones). En nuestro ejemplo:

$$C_T = \frac{C_{1 \times 2} + C_{1 \times 3} + C_{2 \times 3}}{3}$$

$$C_T = \frac{0.94 + 0.88 + 0.88}{3} = 0.90$$

No es conveniente tolerar una confiabilidad menor que 0.85 (ni total ni entre dos 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 total).

En el análisis de contenido una confiabilidad baja puede deberse a que las categorías y/o unidades de análisis no han sido definidas con claridad y precisión, a un deficiente entrenamiento o a inhabilidad de los codificadores. Cuando se obtiene una baja confiabilidad debe detectarse y corregirse el problema. Asimismo, es conveniente calcular la confiabilidad a la mitad de la codificación (con el material codificado) y al finalizar esta.

1. Efectuar la codificación. Lo que implica contar las frecuencias de repetición de las categorías (número de unidades que entran en cada categoría).
2. Variar los datos de las hojas de codificación y obtener totales para cada categoría.
3. Realizar los análisis estadísticos apropiados.

2.5.4 Observación

¿Qué es y para qué sirve la observación?

La observación consiste en el registro sistemático, válido y confiable de comportamiento o conducta manifiesta. Puede utilizarse como instrumento de medición

en muy diversas circunstancias. Havnes (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, eventos masivos (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, es una forma de observación del contenido de comunicaciones. Es por ello que 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 cuentan las 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, sólo 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) 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. 107-108).

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 tensión:** 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 estar rascando, picando, etc.) y la postura en general del sujeto.
3. **Conducta visual del sujeto:** que según lo estipula en esa 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 dicotomas y silencios (evitación), con valor de "0".

La modalidad de frases dicotomas incluye respuestas monosilábicas (sí, no), murmullos, sonidos guturales: los silencios que se interpretan como respuestas dicotomas (respuesta de evitación) son los silencios no naturales en el discurso, en los que expresamente el sujeto se queda 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 principio se pensó que además de medir la conducta verbal en cuanto a su formato, convendría medir 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, debe establecer que en los cuatro indicadores (variables) adoptados para medir la evitación en la interacción se establecieron dimensiones cuya medición fuese dicotómica; es decir, las opciones de respuesta para cada variable únicamente podían ser codificadas bajo la escala de 0 - 1; esto es, evitación o no evitación. Las variables no adquieron más opciones de respuesta por la dificultad de detectar conductas de evitación tanto verbales como no verbales.

Las subcategorías pueden ser escalas de Likert

Las subcategorías pueden ser escalas del tipo Likert, Guttman o diferencial semántica.

Categoría	Subcategorías
Atención del alumno	<ul style="list-style-type: none"> Alta (3) Media (2) Baja (1) Nula (0)
Disciplina del alumno	<ul style="list-style-type: none"> Completa (3) Aceptable (2) Media (1) Indisciplinada (0)
Buen aprendizaje del niño:	<ul style="list-style-type: none"> Muy alta (5) Alta (4) Regular (3) Baja (2) Inexistente (1)

(1) al hablar del desempeño laboral observado en coordinadores de escuelas o asociaciones:

Productividad : (5) (4) (3) (2) (1) (0) : Improductividad

Actitud de servicio : (5) (4) (3) (2) (1) (0) : Actitud de no servicio

Cumplimiento-incumplimiento

Escolera

Seleccionar a los observadores: los observadores son las personas que habrán de codificar la conducta y deben conocer las variables, categorías y subcategorías.

Figurar el medio de observación: la conducta o sus manifestaciones pueden codificarse por distintos medios: puede observarse directamente y ser codificada, puede grabarse en video-cinta y analizarse (con o sin audio, dependiendo del hecho de que se evalúe o no la conducta verbal). En algunos casos el observador se oculta y observa (por ejemplo, a través de un espejo de doble visual). Otras veces participa con los sujetos y codifica. En ciertas ocasiones se codifican manifestaciones de la conducta en la observación es a posteriori (por ejemplo, un estudio para evaluar las condiciones higiénicas de una comunidad o la infraestructura con que cuenta una población). El medio a través del cual se observe depende de la investigación en particular.

Elaborar las hojas de codificación: su formato es el mismo que se presentó en el apartado sobre análisis de contenido. En el caso de Naves y Poplawsky (1984) la hoja de codificación fue la que se muestra en la Figura 9.20.

Proporcionar entrenamiento de codificadores: con 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.

Calcular la confiabilidad de los observadores (intraobservador 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 "codificadores", "codificación", "codificadas"; por "observadores", "observación", "observadas".

Por ejemplo:

$$\text{Confiabilidad} = \frac{\text{Número de unidades de análisis catalogadas correctamente por el observador individual}}{\text{Número total de unidades de análisis}}$$

$$\text{Confiabilidad} = \frac{\text{Número total de acuerdos entre dos parejas entre parejas}}{\text{Número total de unidades de análisis observadas}}$$

Haynes (1978, p. 160) proporciona otra fórmula para calcular la confiabilidad entre observadores o el grado de acuerdo interobservadores (Ao):

$$Ao = \frac{ia}{ia + id}$$

Donde "ia" es el número total de acuerdos entre observadores y "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).

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.

Nombre _____ Grupo _____
 Fecha _____

Categoría	Subcategoría	0'	10"	20"	30"	40"	50"	1'	10"	20"	30"	40"	50"	1'	20"	30"	40"	50"	3'	
Distancia física	Subcategoría																			
	Alcance (0)																			
	Acercamiento (2)																			
	Estático (1)																			
Movimientos corporales	Tensión (0)																			
	Relajación (2)																			
	Ninguna (1)																			
Conducta visual	Al sujeto (1)																			
	A otra parte (0)																			
	F. completas (1)																			
Conducta verbal	F. dicotomas (0)																			
	Situadas (0)																			

Nota: Se agregaron las subcategorías "estático" en la categoría "distancia física" y "ninguno" en la categoría "movimientos corporales".

FIGURA 9.20
 Ejemplo de una hoja de codificación para observar conducta.

1. Llevar a cabo la codificación por observación.
 2. Vaciar los datos de las hojas de codificación y obtener totales para cada categoría.
 3. 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

Ventajas de la observación

Tanto la observación como el análisis de contenido tienen varias ventajas:

1. *Son técnicas de medición no observativas.* En el sentido que el instrumento de medición no "estimula" el comportamiento de los sujetos (las escuelas de actividad y los cuestionarios prefieren "estimular" una respuesta a cada ítem). Los métodos no observativos 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 los 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.³⁶

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).

³⁶ Para conocer la diversidad de estas pruebas y sus aplicaciones se recomienda Anastasi (1982, Thorndike y Hagen (1980), Cronbach (1984) y Nunnally (1978). Son obras clásicas sobre medición y el manejo de pruebas estandarizadas.

En caso de que afirmamos una prueba diseñada en otro contexto, es necesario adaptarla y aplicar pruebas pletas 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 manchas de tinta en tarjetas o láminas blancas numeradas a los sujetos y estos relatan sus asociaciones e interpretaciones en relación a las manchas y el test de *apercpción 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).

Basantes 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 estímulos.

4.3 Sesiones en profundidad

En el 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 este en relación a las variables de la investigación. Pueden realizarse una o varias reuniones. El procedimiento usual es el siguiente.

Definir el tipo de personas que habrán de participar en la sesión o sesiones.

Se define el tipo de personas que habrán de participar en la sesión o sesiones.

Se detectan personas del tipo elegido.

Se invita a estas personas a la sesión o sesiones.

Se organiza 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 (taun las cuestiones más sencillas como servir café y refrescos).

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 dis-

ante" por los participantes de la sesión y debe promover 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, valen diversas 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.

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.

Llevar a cabo la *codificación y análisis correspondiente*.

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 sabor, color, presentación, precio, cualidades del producto, discutir a fondo las propiedades, cantidades y carencias del producto administrando una escala de actitudes o un cuestionario, y hacer preguntas abiertas a cada participante. (También 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 niveles. Por ejemplo, los autores participaron en un estudio para evaluar un programa televisivo que acababa de salir al aire con un nombre y formato nuevos. Además de evaluar una encuesta telefónica se organizaron varias sesiones (algunas con amas de casa, otras con estudiantes, también con trabajadores, empleados de oficina, operativos, 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 30 años, una ama de casa más joven, un empleado de una oficina pública, una secretaria, un profesor, un dependiente de supermercado, un operativo y dos estudiantes).

En las sesiones se profundizó en el formato del programa (historia, manejo de cámaras, duración, manejo de comerciales, sonido, manera de presentar avisados, las 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 acudir a *archivos* que contengan los datos. Por ejemplo, si nuestra hipótesis fue: "La violencia manifiesta en la ciudad de México es mayor que en la ciudad de Caracas", podemos acudir a la alcaldía de algunas ciudades para solicitar datos relacionados con la violencia: número de asaltos, violaciones, robos a casa-habitación, asesinatos, etc. (datos por habitante, distrito y generales). También podemos 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 en que fueron codificados. El intercambio de estos es una práctica común entre investigadores. Además, existen métodos propios de las diferentes ciencias sociales como el *análisis de redes* para evaluar como se manifiesta la comunicación en un sistema social (quien se comunica con quien, quienes distorsionan la información, como fluye la comunicación, quienes son los líderes comunicativos, etc.); *sistemas de medición fisiológica*, *escalas multidimensionales* que miden a los sujetos en varias dimensiones (v.g., el sistema Galileo de J. Woffel y E. J. Fink, 1980), como medir la distancia psicológica entre los conceptos "patria", "madre", "presidente" y "nación", etcétera, tomando en cuenta dimensiones cognitivas y emocionales. Y en fin otros métodos que escapan al nivel introductorio de este libro.

Para el manejo de archivos se recomienda consultar a Webb, Campbell y Schwartz (1966), para el análisis de redes a Rogers y Kincaid (1981) y para escalas multidimensionales a Norton (1980), Woffel y Dares (1980) y, desde luego, las obras clásicas de Torrington (1958) y Rummey, Shephard y Nerove (1972).

9.6.8 Combinación de dos o más instrumentos de recolección de los datos

¿Puede utilizarse más de un tipo de instrumento de medición?

En algunos casos, el investigador utiliza varias formas de medición para tener diferentes enfoques sobre las variables. Por ejemplo, el clima laboral en una organización puede medirse a través de una encuesta utilizando un cuestionario, pero además pueden realizarse varias sesiones en profundidad para solicitar opiniones sobre el clima laboral y los problemas existentes, observarse el comportamiento de los trabajadores y analizar el contenido de sus mensajes dirigidos

de la organización (cartas de sugerencias, letreros pintados en los baños, quejas en sus reuniones, etcétera).

9.7 CÓMO SE CODIFICAN LAS RESPUESTAS A UN INSTRUMENTO DE MEDICIÓN

Hemos dicho que las categorías de un ítem o pregunta y las categorías y subcategorías de contenido u observación deben *codificarse con símbolos o números*. Y deben codificarse porque de lo contrario no puede efectuarse ningún análisis o solo se puede contar el número de respuestas en cada categoría (v.g., 25 contestaron "sí" y 24 respondieron "no"). Pero el investigador se interesa en realizar análisis más allá de un conteo de casos por categoría y la mayoría de los análisis se llevan a cabo por computadora. Para ello es necesario transformar las respuestas en símbolos o valores numéricos. Los datos deben resumirse, codificarse y prepararse para el análisis.

También se comentó que las categorías pueden u o no precodificadas llevar la codificación en el instrumento de medición antes de que este sea aplicado) y que las preguntas abiertas no pueden estar precodificadas. Pero en cualquier caso, una vez que se tienen las respuestas, estas deberán codificarse.

La *codificación* de las respuestas implica *cuatro pasos*:

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. Este 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 () 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 condicionar 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, columnas*.

Suponemos que tenemos una escala Likert con tres ítems (frases):

"La Dirección General de Impuestos Nacionales informa oportunamente sobre lo que le interesa",
"La Dirección General de Impuestos Nacionales informa oportunamente sobre lo que le interesa",
"La Dirección General de Impuestos Nacionales informa oportunamente sobre lo que le interesa"

- (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.

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.

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 la que habrá de analizarse. El apartado "*columna*" dentro del libro de códigos tienen

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
	Frase 2 (servicios)	De acuerdo	4	
	Frase 3 (deshonestidad)	Ni de acuerdo, ni en desacuerdo	3	
		En desacuerdo	2	2
		Muy en desacuerdo	1	
		Muy de acuerdo	5	
		De acuerdo	4	
		Ni de acuerdo, ni en desacuerdo	3	
		En desacuerdo	2	3
		Muy en desacuerdo	1	
		Muy de acuerdo	1	
		De acuerdo	2	
		Ni de acuerdo, ni en desacuerdo	3	
		En desacuerdo	2	4
		Muy en desacuerdo	1	
		Muy de acuerdo	5	

FIGURA 1

Ejemplo de un libro de códigos en el caso del estudio de naves y Poplawsky (1984)¹⁷.

Variable	Categorías	Subcategorías	Códigos	Columnas
Número de sujeto	Sujetos		01 a 30 (hubo 30 sujetos)	1 y 2
Tratamiento experimental	Grupo cultural		1	3
	Grupo socio-psicológico		2	
Conducta de evitación	Distancia física	Alejamiento	0	4
		Acercamiento	2	
		Estanco	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 dicotómicas o silencios	0	
Codificador	LRE		1	8
	LEMM		2	
	CFT		3	

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:

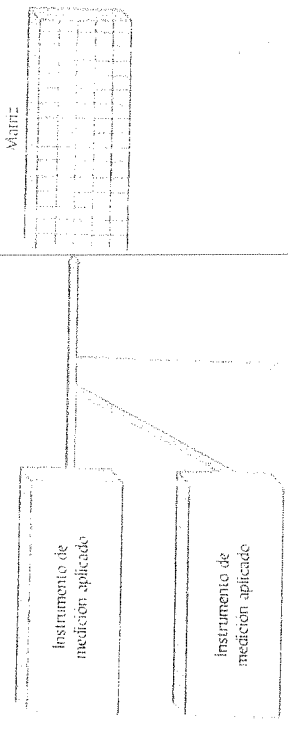
¹⁷ Desde luego, Naves y Poplawsky (1984) para las categorías de "conducta de evitación" abrenían esta codificación cada 10 segundos (cada unidad de análisis), y sumaban el número de 1 (sumas) e lo transformaban en porcentaje. Aquí suponemos que toda la interacción con el deficiente mental puede categorizarse y subcategorizarse.

Matriz de datos

	Columna 1	Columna 2	Columna 3
Caso 1			
Caso 2			
Caso 3			
Caso 4			
Caso K			

Valores de categorías subcategorías

Los resultados del instrumento de medición se transfieren a la matriz por medio del libro de códigos. El proceso puede representarse así:



Libro de códigos

	Columna 1	Columna 2	Columna 3
Caso 1			
Caso 2			
Caso 3			
Caso 4			
Caso K			

En el libro de códigos no puede llegarse a cabo la transferencia. Vamos a ver por qué habríamos aplicado la escala de actitud con esos ítems de la figura 9.21 a cuatro personas, obteniendo los siguientes resultados:

Figura 9.21

A continuación...	El sujeto obtuvo:
"La Dirección General de Impuestos Nacionales informa claramente sobre cómo, dónde y cuándo pagar los impuestos"	
(5) Muy de acuerdo (3) Ni de acuerdo (2) En desacuerdo ni en desacuerdo	4 (de acuerdo)
(4) De acuerdo (1) Muy en desacuerdo	
"Los servicios que presta la Dirección General de Impuestos Nacionales son en general muy buenos"	
(5) Muy de acuerdo (3) Ni de acuerdo (2) En desacuerdo ni en desacuerdo	5 (muy de acuerdo)
(4) De acuerdo (1) Muy en desacuerdo	
"La Dirección General de Impuestos Nacionales se caracteriza por la hostilidad de sus funcionarios"	
(5) Muy de acuerdo (3) Ni de acuerdo (2) En desacuerdo ni en desacuerdo (1) Muy en desacuerdo	3 (ni de acuerdo, ni en desacuerdo)
(4) De acuerdo (1) Muy en desacuerdo	

Persona 1	3 (ni de acuerdo, ni en desacuerdo)
Objetivo respectivamente:	4 (de acuerdo)
	3 (ni de acuerdo, ni en desacuerdo)
Persona 3	4
Objetivo respectivamente:	4
	4
Persona 4	5
Objetivo respectivamente:	4
	3

De acuerdo con el libro de códigos (figura 9.21), tendríamos la matriz de la figura 9.22.

FIGURA 9.22
Ejemplo de matriz de datos para el libro de códigos de la figura 9.21

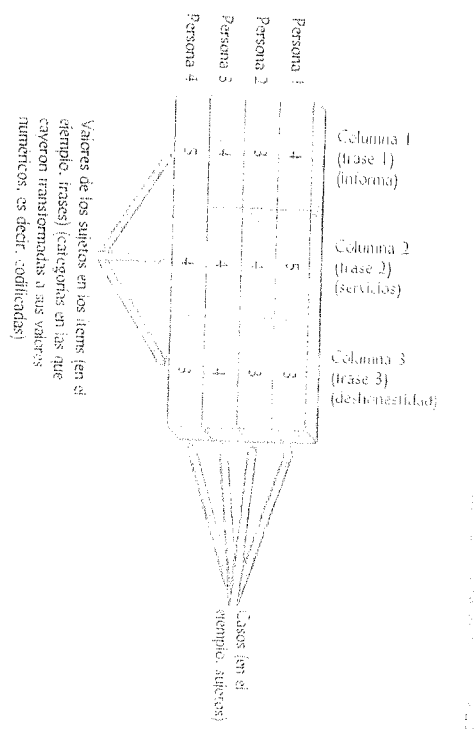


FIGURA 9.24
Ejemplo hipotético de matriz de datos para el libro de códigos de la figura 9.22.

	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
S1	0	1	1	0	0	0	0	1
S2	0	2	1	0	0	0	0	1
S3	0	3	2	2	2	1	1	2
S4	0	4	1	1	0	0	0	2
S5	0	5	2	2	2	1	1	2
S6	0	6	1	0	1	1	0	3
S7	0	7	2	2	2	1	1	3
S8	0	8	2	2	2	1	1	3
S9	0	9	2	1	2	0	1	1
S10	1	0	1	2	1	0	0	2

En el ejemplo de Navés y Poplawsky (figura 9.22), la matriz sería la de la figura 9.24.
El libro de códigos indica a los codificadores qué variable, ítem/categoría, subcategoría van 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 dicotomas 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	0	1	5	
5	0	1	2	2	3	2	3	0	0	0	4	4	
2	1	0	0	0	0	0	0	0	0	0	0	0	
1	1	1	1	1	4	0	2	0	0	0	3	3	
3	0	0	1	2	5	0	2	0	0	0	3	3	
3	1	0	1	2	5	0	1	0	1	0	1	3	
4	0	1	0	2	4	0	1	0	1	0	1	4	
5	1	1	2	2	1	0	0	0	1	0	1	4	
2	0	1	2	2	2	0	2	1	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 libro de códigos y la matriz correspondiente de la figura 9.25.

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
No = 2
No contestó = 3

Contestó incorrectamente = 4 0 Valor perdido por diversas razones = 9

Hasta el momento se han presentado, por razones didácticas, ejemplos resumidos de libros de códigos. Desde luego, un libro de códigos normalmente tiene

FIGURA 9.25
Ejemplo de la variabilidad que pueden tener las columnas libro de códigos.

Variable	Categorías	Códigos	Columna
Número de sujeto	000 a 128	000 a 128	1, 2 y 3
Sexo del respondiente	masculino femenino	1 2	4
Edad del respondiente	00 a 77 No respondió	00 a 77 99	5 y 6
Ingresos del respondiente	Salario mínimo o menos De dos a tres veces el salario mínimo	1 2	7
	De cuatro a cinco veces el salario mínimo	3	
	De seis a ocho veces el salario mínimo	4	
	Nueve a once veces el salario mínimo	5	
	Doce a quince veces el salario mínimo	6	
	16 o más veces el salario mínimo	7	
	No respondió	9	
Tamaño del lugar en que vive (m ²)	0000 a 9998 No respondió	0000 a 9998 9999	8, 9, 10 y 11

Continúa

Matriz de datos

Variable
Nombre de sujeto Sexo Edad Ingresos Tamaño del lugar en que vive (residencia)

Columnas
1 2 3 4 5 6 7 8 9 10 11
Supres

La matriz se llena de acuerdo a los resultados

mas variables o categorías y consecuentemente más columnas tal igual que la matriz de datos. Hay matrices que pueden tener 500 o más columnas. Asimismo, debe recordarse que los renglones son casos (sujeitos, escuelas, series de televisión, etcétera), y a veces es necesario extender las columnas a otro renglón si otros renglones tienen computaciones limitadas a manejar 80 columnas. Esto podrá representarse así:

Col. 1	2	3	4	etc.	col. 80
Col. 31	32	33	34	etc.	col. 160
Col. 161	162	163	164	etc.	col. 240
Col. 241				etc.	col. 320
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

Codificación física

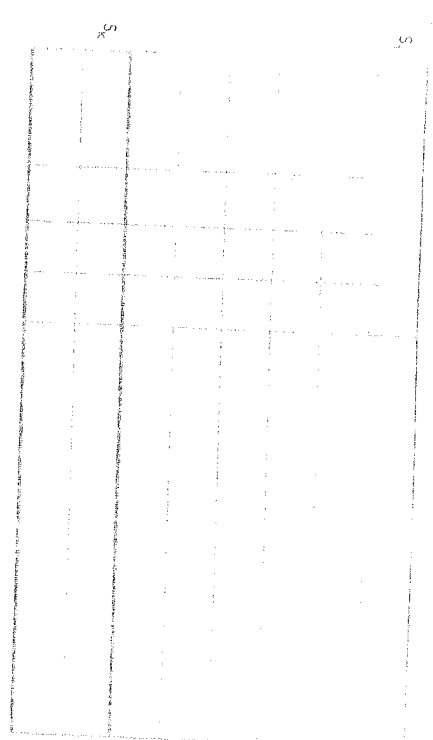
El tercer paso del proceso de codificación es la codificación física de los datos, es decir, llenar la matriz de datos. Esta codificación la efectúan los codificadores, a quienes se les proporciona el libro de códigos. Así, cada codificador va visitando las respuestas en la matriz de datos, de acuerdo con el libro de códigos. El contenido 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 alcanza una hoja de tabulación, se 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.

Aunque actualmente, en la mayoría de las investigaciones ya no se usan las hojas de codificación. Lo común es transferir los datos directamente de los ejemplares del instrumento de medición a la matriz de datos creada como archivo en la computadora. Esta tarea de captura debe ser realizada de manera muy cuidadosa para evitar cometer errores y alterar los resultados.

También existen en la actualidad sistemas de lectura óptica, los cuales pueden leer los datos directamente de los cuestionarios (u otros instrumentos de medición) o de hojas de tabulación. Algunos requieren tintas o cierto tipo de lápiz; otros copian la información mediante un escáner. Desde luego, éstos son sistemas costosos.

Generación de archivos

En cualquier caso, se crea la matriz de datos y se archiva, así se genera un *file* o archivo, el cual debe ser nombrado y contiene los datos codificados en valores numéricos. Hemos generado un archivo con los datos recolectados y codificados. El archivo puede almacenarse en el disco duro, un disquete o un CD-ROM. El proceso se muestra en la figura 9.27.



Fecha _____
 Investigación _____
 Codificación _____
 Hoja No. _____

Columnas (variables)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	2	2	3	4	5	6	1	0											
2	1	1	0	4	4	5	2	0	1											
3	1	1	0	4	4	5	2	1	1											
4	2	1	0	4	4	5	3	1	1											
5	2	1	1	4	5	4	5	1	0											
6	2	1	1	4	5	4	5	1	0											
7	2	1	1	3	5	5	5	0	0											

Ren-
glones
(casos)

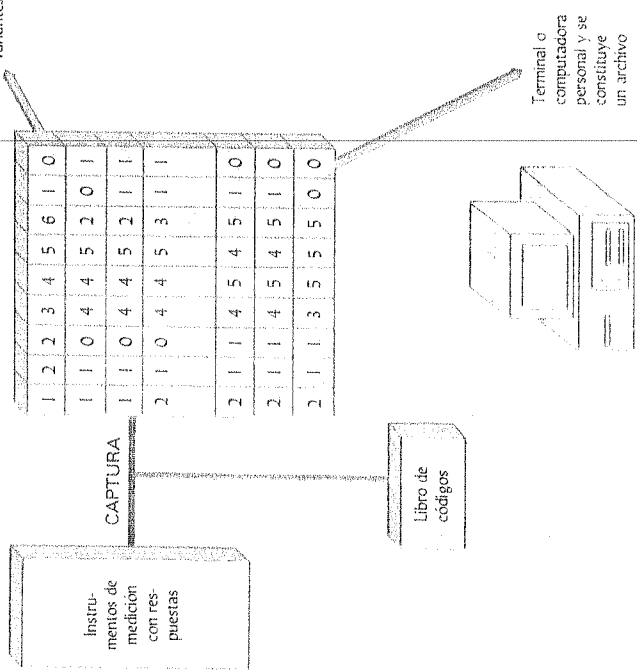


FIGURA 9.27
 Proceso para crear un archivo de datos (archivar la matriz de datos).

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.

Ejemplo

En el caso de personas infectadas por el VIH tras recibir sangre o sus derivados provenientes del laboratorio "Transfusiones y Hematología, S.A.", los datos de la evolución de los pacientes se recolectaron a través del tiempo mediante la aplicación de análisis de laboratorio, reconocimientos médicos, presentación de sintomatología. De esta forma se cuantificaron datos (el más relevante, ver que de los 49 infectados

<ul style="list-style-type: none"> a) 25 eran trabajadores de Pernes y 24 familiares de los mismos. b) La edad de los afectados oscilaba de los 2 años hasta los 74. c) 25 eran hombres y 24 mujeres. d) Que en grado I tenían 0 pacientes; en grado II, 6; en grado III, 16; y en grado IV, 9, además de los 18 fallecidos. e) La tasa de transisiones con VIII fue de 1,7 por cada 100. <p>A través de estos datos no se puede asegurar la evolución de los pacientes infectados que no entraron en la muestra de los 49 porque, aunque los instrumentos para medir los datos que se acaban de presentar son los mismos que se utilizan en otras instituciones, no permiten la predictibilidad respecto a pacientes no analizados directamente. Lo que sí se puede asegurar es que los datos arriba descritos son ciertos, puesto que se obtuvieron de mecanismos de medición aceptados por la comunidad médica como válidos.</p>	<ul style="list-style-type: none"> • No hay medición perfecta, pero el error de medición debe reducirse a límites tolerables. • La confiabilidad se determina calculando un coeficiente de confiabilidad. • Los coeficientes de confiabilidad varían entre 0 y 1 (0 = nula confiabilidad, 1 = total confiabilidad). 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 alfa de Cronbach y el coeficiente KR-20. • La validez de contenido se obtiene comparando el universo de items contra los items presentes en el instrumento de medición. • La validez de criterio se obtiene comparando los resultados de aplicar el instrumento de medición contra los resultados de un criterio externo. • La validez de constructo se puede determinar mediante el análisis de factores. • Los pasos genéricos para construir un instrumento de medición son: <ul style="list-style-type: none"> — Listar las variables a medir. — Revisar sus definiciones conceptuales y operacionales. — Elegir uno ya desarrollado o construir uno propio.
--	--

RESUMEN

- 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.
- Medir es el proceso de vincular conceptos abstractos con indicadores empíricos, mediante clasificación y/o cuantificación.
- En toda investigación medimos las variables contenidas en la hipótesis.
- Un instrumento de medición debe cubrir dos requisitos: confiabilidad y validez.
- La confiabilidad se refiere al grado en que la aplicación repetida de un instrumento de medición al mismo sujeto u objeto, produciendo iguales resultados. La validez refiere al grado en que un instrumento de medición mide realmente (las variables) que pretende medir.
- 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.
- 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.

- Indicar niveles de medición de las variables (nominal, ordinal, por intervalos y de razón).
 - Indicar como se habrán de codificar los datos.
 - Aplicar prueba piloto.
 - Construir versión definitiva.
- 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.
- Las respuestas se codifican. 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
- Coefficiente de confiabilidad
- Niveles de medición
- Medida de estabilidad
- Método de formas alternas
- Método de mitades parciales
- Coefficiente alfa de Cronbach
- Evidencia relacionada con el contenido
- Evidencia relacionada con el criterio
- Evidencia relacionada con el constructo
- Escalas de actitudes
- Escala Likert
- Diferencial semántico
- Esquema 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 1) 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 items para medir dicha actitud e indique cómo se calificaría la escala (total 15 items positivos y 5 negativos). Finalmente indique la dimensión que cada item pretende medir de dicha actitud (credibilidad, presencia física, etcétera).

1. Construya un cuestionario para medir lo que usted considere conveniente (con preguntas demográficas y, por lo menos, 10 preguntas más), añada a 20 empujados 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. Planee una sesión en profundidad (indique objetivos, procedimientos, sujetos tipo, agenda, etc.)

BIBLIOGRAFÍA SUGERIDA

- ANASTASI, A. (1982). *Psychological testing*. Nueva York, NY: MacMillan Publishing Co., Inc. 5a. edición.
- RABBIE, 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 Hill, CA: Sage Publications, Inc. Series: Quantitative Applications in the Social Sciences, Vol. 17.
- CRONBACH, L.J. (1984). *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. 3a. edición.
- MUNNALLY, J.C. (1970). *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). *Experimental measures: nonreactive research in the social sciences*. USA: Rand McNally College Publishing Co.
- WIERSMA, W. (1986). *Research methods in education*. Newton, Mass: Allyn and Bacon, Inc. 4a. edición, capítulo 1 ("Measurement and data collection").

Ejemplo

LA TELEVISIÓN Y EL NIÑO

Cuestionarios que se suministraron a los niños del Distrito Federal para el ejemplo de "La televisión y el niño" (se incluye parcialmente en versión precodificada)

¿Cada cuánto veces 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)

¿Cada cuánto vas al cine?

- _____ una vez por semana (4)
 _____ una cada dos semanas (3)
 _____ una vez al mes (2)
 _____ una vez cada 2 o 3 meses (1)
 _____ casi nunca (0)

¿Cuanto tiempo vas al radio cada día?

- _____ no oigo el radio (0)
 _____ menos de 1 hora al día (1)
 _____ de 1 a 2 horas al día (2)
 _____ de 2 a 3 horas al día (3)
 _____ mas de 3 horas al día (4)

4. Durante la semana pasada, ¿cuantos cuentos de puntos has leído?

_____ (numero)

5. Durante la semana pasada, ¿cuantas revistas has leído, sin tomar en cuenta los cuentos de monitos?

_____ (numero)

6. Sin contar los libros que hayas leído para la escuela, ¿cuantos libros has leído durante el mes pasado?

_____ (numero)

¿Ves tele todos los días?

- Si _____ (1)
 No _____ (2)

En los días que vas a la escuela, ¿cuanto tiempo ves televisión?

- _____ no ves televisión (0)
 _____ menos de 1 hora (1)
 _____ 1 o 2 horas (2)
 _____ 3 o 4 horas (3)
 _____ 5 horas o mas (4)

¿Como cuanto tiempo ves televisión los sábados?

- _____ no ves televisión (0)
 _____ menos de 1 hora (1)
 _____ 1 o 2 horas (2)
 _____ 3 o 4 horas (3)
 _____ 5 horas o mas (4)

10. ¿Cuánta televisión ves los domingos?
no veo televisión (0)
menos de 1 hora (1)
1 o 2 horas (2)
3 o 4 horas (3)
5 horas o más (4)

11. ¿Qué programas de televisión viste ayer?

12. ¿Cuándo ves más televisión?
entre semana (1)
los fines de semana (2)
siempre (3)

13. ¿Cuándo te gusta ver más televisión?
en la tarde (1)
en la noche (2)

14. ¿Qué prefieres hacer cuando no estás en la escuela?
estar con tus papás (1)
jugar (2)
leer (3)
salir a la calle (4)
ver la televisión (5)

15. ¿Que haces cuando vas a ver televisión?
prendo la televisión para ver lo que hay (1)
veo la guía de programas o el periódico, para ver que programas hay en la televisión (2)
ya me sé de memoria lo que hay en la televisión (3)
prendo la televisión y veo lo que sea (4)

16. Mientras estás viendo la televisión, ¿qué es lo que haces generalmente?
como, juego, dibujo o hago cualquier cosa (1)
hago la tarea (2)
veo varios programas a la vez, cambiando de canal (3)

17. De la siguiente lista de programas marca con una palomita los que tú ves, y además qué tanto te gustan.

	Me gusta mucho (1)	Me gusta un poco (2)	No lo veo (0) (3)
PorKy			
Mi maridano favorito			
Mundo de juguete			
Varietades vergel			
El hombre nuclear			
Noticario Domneq			
Operación convivencia			
Futbol			
Películas			
Los Rapiedra			
Hechizada			
Ven conmigo			
La criada bien criada			
Kojack			
En punto			
Plaza Sesamo			
El Oso Ruperto			
Locos Adams			
Una muchacha llamada Milagros			
Los polivoces			
Viaje al fondo del mar			
24 horas			
Inverso 5			
Clasicos infantiles			
Nani y el profesor			
Batata 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)
- ___ sirvienta (7)

20. Marca con una palomita, si estás de acuerdo o no, con las siguientes frases.

Mi papá o mi mamá me regañan cuando veo mucha televisión.

- Sí ___ No ___
- (1) (0)

A veces me castigan sin ver televisión

- Sí ___ No ___
- (1) (0)

Mi papá o mi mamá me prohíben ver algunos programas de televisión.

- Sí ___ No ___
- (1) (0)

Yo me voy a dormir a la hora que quiero.

- Sí ___ 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. Veo 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)