

Se explicará esto con un ejemplo coloquial. Si fuera a apostar en las carreras de caballos y tuviera 95% de probabilidades de atinarle al ganador, contra sólo 5% de perder, ¿apostarías? Obviamente sí, siempre y cuando le aseguraran ese 95% en favor. O bien, si le dieran 95 boletos de 100 para la rifa de un automóvil, ¿sentirías confianza en que va a estrenar vehículo? Por supuesto que sí. No tendría la certeza total; ésta no existe en el universo, al menos para los seres humanos.

Pues bien, algo similar hace el investigador al definir un posible nivel de error en la representatividad de su muestra. Los niveles de error más comunes que suelen fijarse en la investigación son de 5 y 1% (en ciencias sociales el más usual es el primero).

El porcentaje estimado de la muestra es la probabilidad de ocurrencia del fenómeno (representatividad de la muestra *versus* no representatividad, la cual se estima sobre marcos de muestreo previos o se define, la certeza total siempre es igual a uno, las posibilidades a partir de esto son “*p*” de que sí ocurra y “*q*” de que no ocurra ($p + q = 1$). Cuando no tenemos marcos de muestreo previo, usamos un porcentaje estimado de 50% (que es la opción por “default” que nos brinda STATS®, es decir, asumimos que “*p*” y “*q*” serán de 50%, y que resulta lo más común, particularmente cuando seleccionamos por vez primera una muestra en una población).

Finalmente, el nivel deseado de confianza es el complemento del error máximo aceptable (porcentaje de “acertar en la representatividad de la muestra”). Si el error elegido fue de 5%, el nivel deseado de confianza será de 95%. Una vez más los niveles más comunes son de 95 y 99%. Por default, STATS® coloca el primero.

Ya con todos los “campos” llenos, con sólo presionar el botón de “calcular”, se obtiene el tamaño de muestra apropiado para el universo. En el ejemplo podría ser:

¿Tamaño del universo?: 2200
 ¿Error máximo aceptable?: 5%
 ¿Porcentaje estimado de la muestra?: 50%
 ¿Nivel deseado de confianza?: 95

El resultado que nos proporciona STATS® es:

Tamaño de la muestra: 327.1776. Redondeando, necesitamos que nuestra muestra esté conformada por 327 empresas para tener representadas a las 2200 de la ciudad.



EJEMPLO

Problema de investigación:

Supongamos que el gobierno de un estado, provincia o departamento ha emitido una ley que impide (prohibición expresa) a las estaciones de radio transmitir comerciales que utilicen un lenguaje procaz (groserías, malas palabras). Dicho gobierno nos solicita analizar en qué medida los anuncios radiofónicos transmitidos en el estado utilizan en su contenido este lenguaje, digamos, durante el último mes.

Población (N):

Comerciales transmitidos por las estaciones radiofónicas del estado durante el último mes.

Tamaño de muestra (n):

Lo primero es determinar o conocer *N* (recordemos que significa población o universo). En este caso $N = 20000$ (20 mil comerciales transmitidos). Lo segundo es establecer el error máximo aceptable, el porcentaje estimado de la muestra y el nivel de confianza.

Tecleanos los datos que STATS® nos pide:

Tamaño de la población: 20000
 Error máximo aceptable: 5%
 Porcentaje estimado de la muestra: 50%
 Nivel de confianza: 95%

De manera automática, el programa nos calcula el tamaño de muestra necesario o requerido: $n = 376.9386$ (cerrando o aproximando: 377), que es el número de comerciales radiofónicos que necesitamos para representar al universo de 20 000, con un error de 0.05 (5%) y un nivel de confianza de 95%.

Si cambiamos el nivel de error tolerado y el nivel de confianza (1% de error y 99% de confianza), el tamaño de la muestra será mucho mayor, en este caso de 9 083.5153 comerciales.

Como puede apreciarse, el tamaño de la muestra es sensible al error y nivel de confianza que definamos. A menor error y mayor nivel de confianza, mayor tamaño de muestra requerido para representar a la población o universo.

EJEMPLO

Problema de investigación:

Analizar la motivación intrínseca que tienen los empleados de la cadena de restaurantes “Lucy y Laura Bunny”.

Población:

$N = 600$ empleados (cocineros, meseros, ayudantes, etcétera).

Tamaño de muestra:

Con un error de 5% y un nivel de confianza de 95%, el tamaño requerido para que la muestra sea representativa es de 234 empleados.

Conforme disminuye el tamaño de la población aumenta la proporción de casos que necesitamos en la muestra.

Previamente se señaló que para obtener una muestra probabilística eran necesarios dos procedimientos, el primero es el que acabamos de mencionar: calcular un tamaño de muestra que sea representativo de la población. El segundo consiste en seleccionar los elementos muestrales de manera que al inicio todos tengan la misma posibilidad de ser elegidos. Es decir, cómo y de dónde vamos a elegir los casos. Esto se comentará más adelante.



A los ejemplos de las muestras obtenidas por STATS® se les conoce como muestras aleatorias simples (MAS). Su característica esencial, como ya se mencionó, es que todos los casos del universo tienen al inicio la misma probabilidad de ser seleccionados.

Muestra probabilística estratificada

Muestra probabilística estratificada Muestreo en el que la población se divide en segmentos y se selecciona una muestra para cada segmento.

En ocasiones el interés del investigador es comparar sus resultados entre segmentos, grupos o nichos de la población, porque así lo señala el planteamiento del problema. Por ejemplo, efectuar comparaciones por género (entre hombres y mujeres), si la selección de la muestra es aleatoria, tendremos unidades o elementos de ambos géneros, no

hay problema, la muestra reflejará a la población.

Pero a veces, nos interesan grupos que constituyen minorías de la población o universo y entonces si la muestra es aleatoria simple, resultará muy difícil determinar qué elementos o casos de tales grupos serán seleccionados. Imaginemos que nos interesan personas de todas las religiones para contrastar ciertos datos, pero en la ciudad donde se efectuará el estudio la mayoría es —por ejemplo— predominantemente católica. Con MAS es casi seguro que no elijamos individuos de diversas religiones o sólo unos cuantos. No podríamos efectuar las comparaciones. Quizá tengamos 300 católicos y dos o tres de otras religiones. Entonces es cuando preferimos obtener una *muestra probabilística estratificada* (el nombre nos dice que será probabilística y que se considerarán segmentos o grupos de la población, o lo que es igual: estratos).

Ejemplos de estratos en la variable religión serían: católicos, cristianos, protestantes, judíos, mahometanos, budistas, etc. Y de la variable grado o nivel de estudios: preescolar, primaria, secundaria, bachillerato, universidad (o equivalente) y posgrado.

Los ejemplos anteriores para ilustrar el uso de STATS® corresponden a muestras probabilísticas simples. Ahora supongamos que pretendemos realizar un estudio con directores de recursos humanos para determinar su ideología y políticas respecto a cómo tratan a los colaboradores de sus empresas. Imaginemos que nuestro universo es de 1 176 organizaciones con directores de recursos humanos. Usando STATS® o mediante fórmulas, determinamos que el tamaño de la muestra necesaria para representar a la población sería de $n = 298$ directivos. Pero supongamos que la situación se complica y que debemos estratificar esta n con la finalidad de que los elementos muestrales o las unidades de análisis posean un determinado atributo. En nuestro ejemplo, este atributo podría ser el giro de la empresa. Es decir, cuando no basta que cada uno de los elementos muestrales tengan la misma probabilidad de ser escogidos, sino que además es necesario segmentar la muestra en relación con estratos o categorías que se presentan en la población, y que además son relevantes para los objetivos del estudio, se diseña una muestra probabilística estratificada. Lo que aquí se hace es dividir a la población en subpoblaciones o estratos, y se selecciona una muestra para cada estrato.

La estratificación aumenta la precisión de la muestra e implica el uso deliberado de diferentes tamaños de muestra para cada estrato, a fin de lograr reducir la varianza de cada unidad de la media muestral (Kalton y Heeringa, 2003). Kish (1995) afirma que, en un número determinado de elementos muestrales $n = \sum nb$, la varianza de la media muestral \bar{y} puede reducirse al mínimo, si el tamaño de la muestra para cada estrato es proporcional a la desviación estándar dentro del estrato.

Esto es,

$$\sum fh = \frac{n}{N} = ksh$$

En donde la muestra n será igual a la suma de los elementos muestrales nb . Es decir, el tamaño de n y la varianza de \bar{y} pueden minimizarse, si calculamos “submuestras” proporcionales a la desviación estándar de cada estrato. Esto es:

$$fh = \frac{nb}{Nh} = ksh$$

En donde nb y Nh son muestra y población de cada estrato, y sh es la desviación estándar de cada elemento en un determinado estrato. Entonces tenemos que:

$$ksh = \frac{n}{N}$$

Siguiendo con nuestro ejemplo, la población es de 1 176 directores de recursos humanos y el tamaño de muestra es $n = 298$. ¿Qué muestra necesitaremos para cada estrato?

$$ksh = \frac{n}{N} = \frac{298}{1\,176} = 0.2534$$

De manera que el total de la subpoblación se multiplicará por esta fracción constante para obtener el tamaño de la muestra para el estrato. Al sustituirse, tenemos que:

$$(Nh)(fh) = nb \text{ (véase tabla 8.2)}$$



▲ **Tabla 8.2** Muestra probabilística estratificada de directores de empresa

Estrato por giro	Directores de recursos humanos del giro	Total población (fh) = 0.2534 Nh (fh) = nh	Muestra
1	Extractivo y siderúrgico	53	13
2	Metal-mecánicas	109	28
3	Alimentos, bebidas y tabaco	215	55
4	Papel y artes gráficas	87	22
5	Textiles	98	25
6	Eléctricas y electrónicas	110	28
7	Automotriz	81	20
8	Químico-farmacéutica	221	56
9	Otras empresas de transformación	151	38
10	Comerciales	51	13
		$N = 1176$	$n = 298$

Por ejemplo:
 $Nh = 53$ directores de empresas extractivas corresponden a la población total de este giro.
 $fh = 0.2534$ es la fracción constante.
 $nh = 13$ es el número redondeado de directores de empresa del giro extractivo y siderúrgico que tendrá que entrevistarse.

Muestreo probabilístico por racimos

En algunos casos en que el investigador se ve limitado por recursos financieros, por tiempo, por distancias geográficas o por una combinación de éstos y otros obstáculos, se recurre al *muestreo por racimos* o *clusters*. En este tipo de muestreo se reducen costos, tiempo y energía, al considerar que muchas veces las unidades de análisis se encuentran encapsuladas o encerradas en determinados lugares físicos o geográficos, a los que se denomina **racimos**. Para dar algunos ejemplos tenemos la tabla 8.3. En la primera columna se encuentran unidades de análisis que frecuentemente vamos a estudiar. En la segunda columna, sugerimos posibles racimos donde se encuentran dichos elementos.

Racimos Son sinónimos de *clusters* o conglomerados.

▲ **Tabla 8.3** Ejemplo de racimos o *clusters*

Unidad de análisis	Posibles racimos
Adolescentes	Preparatorias
Obreros	Industrias o fábricas
Amas de casa	Mercados/supermercados/ centros comerciales
Niños	Colegios

Muestra probabilística por racimos Muestreo en el que las unidades de análisis se encuentran encapsuladas en determinados lugares físicos.

Muestrear por racimos implica diferenciar entre la unidad de análisis y la unidad muestral. La unidad de análisis indica quiénes van a ser medidos, o sea, los participantes o casos a quienes en última instancia vamos a aplicar el instrumento de medición. La unidad muestral (en este tipo de muestra) se refiere al racimo por medio del cual se logra el acceso a la unidad de análisis. El muestreo por racimos supone una selección en dos o más etapas, todas con procedimientos probabilísticos. En la primera, se seleccionan los racimos, siguiendo los pasos ya señalados de una muestra probabilística simple o estratificada. En las fases subsecuentes y dentro de estos racimos, se seleccionan los casos que van a medirse. Para ello se hace una selección que asegure que todos los elementos del racimo tienen la misma probabilidad de ser elegidos.

Por ejemplo, en una muestra nacional de ciudadanos de un país por *clusters* o racimos, podríamos primero elegir al azar una muestra de estados, provincias o departamentos (primera etapa); luego, cada estado o provincia se convierte en un universo y se seleccionan al azar municipios (segunda etapa); posteriormente, cada municipio se considera un universo o población y se eligen al azar comunidades o colonias (tercera etapa); a su vez, cada una de éstas se concibe como universo y de nuevo, al azar, se eligen manzanas o cuadras (cuarta etapa); finalmente se escogen al azar viviendas u hogares e individuos (quinta etapa).

En ocasiones se combinan tipos de muestreo, por ejemplo: una muestra probabilística estratificada y por racimos, pero siempre se utiliza una selección aleatoria que garantiza que al inicio del procedimiento todos los elementos de la población tienen la misma probabilidad de ser elegidos para integrar la muestra. En el CD anexo: Material complementario → Documentos → Documento “Cálculo del tamaño de muestra y otros procedimientos por fórmulas” se proporciona un ejemplo que comprende varios de los procedimientos descritos hasta ahora y que ilustra la manera como frecuentemente se hace una muestra probabilística en varias etapas por conglomerados o racimos.



¿Cómo se lleva a cabo el procedimiento de selección de la muestra?

Cuando iniciamos nuestra exposición sobre la muestra probabilística, señalamos que los tipos de muestra dependen de dos cosas: del tamaño de la muestra y del procedimiento de selección.

De lo primero hemos hablado con detalle, de lo segundo trataremos ahora. Se determina el tamaño de la muestra n , pero ¿cómo seleccionar los elementos muestrales? (ya sean casos o racimos). Las unidades de análisis o los elementos muestrales se eligen siempre aleatoriamente para asegurarnos de que cada elemento tenga la misma probabilidad de ser elegido. Se utilizan básicamente tres procedimientos de selección, de los cuales a continuación se comentan dos y el tercero se presenta en el CD anexo: Material complementario → Documentos → Documento “Cálculo del tamaño de muestra y otros procedimientos por fórmulas”.



Tómbola

Muy simple pero muy rápido, consiste en numerar todos los elementos muestrales de la población, del uno al número N . Después se hacen fichas o papeles, uno por cada elemento, se revuelven en una caja y se van sacando n número de fichas, según el tamaño de la muestra. Los números elegidos al azar conformarán la muestra.

Cuando nuestro muestreo es estratificado, se sigue el procedimiento anterior, pero por cada estrato. Por ejemplo, en la tabla 8.2, tenemos que, de una población $N = 53$ empresas extractivas y siderúrgicas, se necesita una muestra $n = 13$ de directivos de recursos humanos de tales empresas. En una lista se numeran cada una de estas organizaciones. En fichas aparte se sortea cada uno de los 53 números, hasta obtener los 13 necesarios (pueden ser las 13 primeras fichas que se extraigan). Los números obtenidos se verifican con los nombres y las direcciones de nuestra lista, para precisar los directivos que serán participantes del estudio.

Números *random* o números aleatorios

Éste es el procedimiento que se encuentra en el CD anexo: Documento “Cálculo del tamaño de muestra y otros procedimientos por fórmulas”.



STATS®

Una excelente alternativa para generar números aleatorios se encuentra en el programa STATS®, que contiene un subprograma para ello y evita el uso de la tabla de números aleatorios. Es hasta ahora la mejor forma que hemos encontrado para hacerlo.

El programa nos pide que le indiquemos ¿cuántos números aleatorios? (requerimos), entonces tecleamos el tamaño de muestra; en el CD elegimos la opción: “Establecer límite superior e inferior”

STATS®



y nos solicita que establezcamos el límite inferior (que siempre será uno, el primer caso de la población, pues la muestra se extrae de ésta) y el límite superior (el último número de la población, que es el tamaño de la población). Y tecleamos “Calcular” y nos genera automáticamente los números. Vemos contra nuestro listado a quién o a qué corresponde cada número y estos números son los casos que pasarían a integrar la muestra.

Veámoslo con un ejemplo. Imaginemos que una investigadora busca conocer en una escuela o facultad de una universidad quiénes son el joven y la joven más populares. Entonces decide realizar una encuesta, para lo cual debe obtener una muestra. Supongamos que la escuela tiene una población de 1000 alumnos y alumnas. Si obtuviera una muestra aleatoria simple, su procedimiento sería el que se muestra en la figura 8.5.

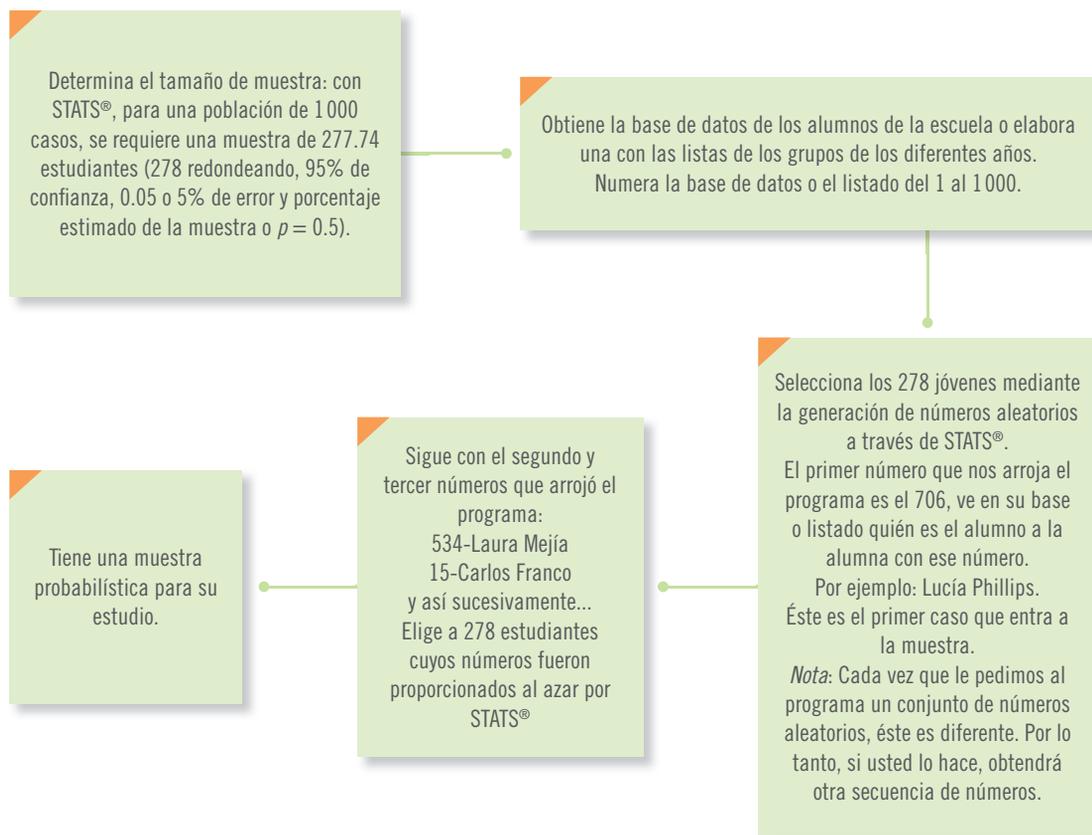


Figura 8.5 Ejemplo del procedimiento para elegir los casos de una muestra aleatoria simple usando STATS®.

Con estratos o conglomerados repetimos el procedimiento para cada uno.

Selección sistemática de elementos muestrales

Este procedimiento de selección es muy útil e implica elegir dentro de una población N un número n de elementos a partir de un intervalo K . Este último (K) es un intervalo que se va a determinar por el tamaño de la población y el tamaño de la muestra. De manera que tenemos que $K = N/n$, en donde K = un intervalo de selección sistemática, N = la población y n = la muestra.

Ilustremos los conceptos anteriores con un ejemplo. Supongamos que se quiere hacer un estudio que pretende medir la calidad de la atención en los servicios proporcionados por los médicos y las enfermeras de un hospital. Para tal efecto consideremos que los investigadores consiguen grabaciones

de todos los servicios efectuados durante un periodo determinado.⁵ Supongamos que se hayan filmado 1 548 servicios (N). Con este dato se procede a determinar qué número de servicios necesitamos analizar para generalizar a toda la población nuestros resultados. Con STATS® determinamos que se necesitan 307.9 (308) servicios para evaluar (con un error máximo de 5%, nivel de confianza de 95% y un porcentaje estimado de 50% para la muestra [$p = 0.5$]).

Si necesitamos una muestra de $n = 308$ episodios de servicio filmados, se utiliza para la selección el intervalo K , donde:

$$K = \frac{N}{n} = \frac{1\,548}{308} = 5.0259, \text{ redondeado} = 5$$

El intervalo $1/K = 5$ indica que cada quinto servicio $1/K$ se seleccionará hasta completar $n = 308$.

La selección sistemática de elementos muestrales $1/K$ se puede utilizar al elegir los elementos de n para cada estrato o para cada racimo. La regla de probabilidad, según la cual cada elemento de la población debe tener idéntica probabilidad de ser elegido, se cumple al empezar la selección de $1/K$ al azar. Siguiendo nuestro ejemplo, no comenzamos a elegir de los 1 548 episodios, el 1, 6, 11, 16..., sino que procuramos que el inicio sea determinado por el azar. Así, en este caso, podemos tirar unos dados y si en sus caras muestran 1, 6, 9, iniciaremos en el servicio 169, y seguiremos 174, 179, 184, 189... $1/K$... y volveremos a empezar por los primeros si es necesario. Este procedimiento de selección es poco complicado y tiene varias ventajas: cualquier tipo de estratos en una población X se verán reflejados en la muestra. Asimismo, la selección sistemática logra una muestra proporcionada, ya que, por ejemplo, tenemos que el procedimiento de selección $1/K$ nos dará una muestra con nombres que inician con las letras del abecedario, en forma proporcional a la letra inicial de los nombres de la población.

Listados y otros marcos muestrales

Las *muestras probabilísticas* requieren la determinación del tamaño de la muestra y de un proceso de selección aleatoria que asegure que todos los elementos de la población tengan la misma probabilidad de ser elegidos. Todo esto lo hemos visto, aunque nos falta exponer sobre algo esencial que precede a la selección de una muestra: el **marco muestral**. Éste constituye un marco de referencia que nos permita identificar físicamente los elementos de la población, la posibilidad de enumerarlos y, por ende, de proceder a la selección de los elementos muestrales (los casos de la muestra). Normalmente se trata de un listado existente o una lista que es necesario confeccionar *ad hoc*, con los casos de la población.

Marco muestral Es un marco de referencia que nos permite identificar físicamente los elementos de la población, así como la posibilidad de enumerarlos y seleccionar los elementos muestrales.

Los listados existentes sobre una población son variados: guías telefónicas, listas de miembros de las asociaciones, directorios especializados, listas oficiales de escuelas de la zona, bases de datos de los alumnos de una universidad o de los clientes de una empresa, registros médicos, catastros, nóminas de una organización, etc. En todo caso hay que tener en cuenta lo completo de una lista, su exactitud, su veracidad, su calidad y su nivel de cobertura en relación con el problema a investigar y la población que va a medirse, ya que todos estos aspectos influyen en la selección de la muestra.

Por ejemplo, para algunas encuestas se considera que el directorio telefónico (o guía telefónica) es muy útil. Sin embargo, hay que tomar en cuenta que muchos números no aparecerán porque son privados o porque hay hogares que no tienen teléfono. La lista de socios de una agrupación como la Cámara Nacional de la Industria de la Transformación (México), la Confederación Española de la

⁵ Se sabe que el número de servicios en un hospital es muy variable y depende de diversos factores, como el número de camas, de médicos y paramédicos; el tipo y nivel de atención (desde consultas simples hasta cirugía compleja), la época, el número de habitantes en la zona donde se encuentra ubicado o el número de derechohabientes, etc. El ejemplo trata de ser simple para que sea entendido por lectores de diversos campos.

Pequeña y Mediana Empresa, la Asociación Dominicana de Exportadores o la Cámara Nacional de Comercio, Servicios y Turismo de Chile; nos serviría si el propósito del estudio fuera, por ejemplo, conocer la opinión de los asociados con respecto a una medida gubernamental. Pero si el objetivo de la investigación es analizar la opinión del sector patronal o empresarial del país, el listado de una sola asociación no sería adecuado por varias razones: hay otras sociedades patronales,⁶ las asociaciones son voluntarias y no todo patrón o empresa pertenece a ellas. Lo correcto, en esta situación, sería construir una nueva base de datos, fundamentada en los listados existentes de las asociaciones patronales, eliminando de dicha lista los casos duplicados, suponiendo que una o más empresas pudieran pertenecer a dos agrupaciones al mismo tiempo.

Hay listas que proporcionan una gran ayuda al investigador. Por ejemplo: bases de datos locales especializadas en las empresas, como Industridata en México;⁷ bases de datos internacionales de naturaleza empresarial como Kompass; directorios por calles o los programas computacionales que tienen a nivel regional o mundial tales directorios; guías de medios de comunicación (que enlistan casas productoras, estaciones de radio y televisión, periódicos y revistas). Este tipo de marcos de referencia construidos por profesionales resultan convenientes para el investigador, pues representan una compilación (de personas, empresas, instituciones, etc.), resultado de horas de trabajo e inversión de recursos. También en internet descubriremos muchos directorios, a los cuales podemos acceder mediante un motor de búsqueda. Recomendamos, pues, utilizarlos cuando sea pertinente, tomando en cuenta las consideraciones que estos directorios o bases de información hacen en su introducción y que revelan el año a que pertenecen los datos, cómo se obtuvieron éstos (exhaustivamente, por cuestionarios, por voluntarios) y muy importante, quiénes y por qué quedan excluidos del directorio.

Con frecuencia es necesario construir listas *ad hoc*, a partir de las cuales se elegirán los elementos que constituirán las unidades de análisis en una determinada investigación. Por ejemplo, en la investigación planteada: La televisión y el niño, se establecería una muestra probabilística estratificada por racimos, donde en una primera etapa se seleccionarían escuelas para, en última instancia, llegar a los niños. Pues bien, para tal efecto se podría obtener una base de datos de las escuelas primarias de la Ciudad de México en la Secretaría de Educación Pública. Cada escuela tendría un código identificable por medio del cual se eliminarían las escuelas para niños atípicos. Este listado contiene además información sobre cada escuela, su ubicación y su régimen de propiedad (pública o privada).

Con ayuda de otro estudio (Fernández Collado *et al.*, 1998) que catalogó en diferentes estratos socioeconómicos a las colonias de la Ciudad de México, con base en el ingreso promedio de la zona, se elaboraron ocho listas:

1. escuelas públicas clase A;
2. escuelas privadas clase A;
3. escuelas públicas clase B;
4. escuelas privadas clase B;
5. escuelas públicas clase C;
6. escuelas privadas clase C;
7. escuelas públicas clase D;
8. escuelas privadas clase D.

Cada lista representaría un estrato de la población y de cada una de ellas se seleccionaría una muestra de escuelas. A, B, C, D, que representan niveles socioeconómicos. Y después, de cada escuela se elegirían los niños para conformar la muestra final.

⁶ En México la Canacindra representa sólo al sector de la industria de la transformación, en España la Cepyme no agrupa a grandes consorcios empresariales, en República Dominicana la Adoexpo no es la única asociación del Consejo Nacional de la Empresa Privada y en Chile la CNC no incluye a la industria de la construcción y la minería, por ejemplo.

⁷ Directorio que permite consultar información de empresas por giro de actividad: industriales, comerciales, de servicio y constructoras, así como el número de personas empleadas. La base de datos clasifica a dichas compañías en: empresas AAA, con más de 500 personas empleadas; empresas AA que tienen entre 251 y 500 personas empleadas; empresas A, entre 151 y 250 personas, y empresas B, entre 100 y 150 personas empleadas.

No siempre existen listas que permitan identificar a nuestra población. Entonces, será necesario recurrir a otros marcos de referencia que contengan descripciones del material, las organizaciones o los casos que serán seleccionados como unidades de análisis. Ejemplos de algunos de estos marcos de referencia son los archivos, los mapas y los archivos electrónicos de periódicos en la web. De cada una de estas instancias daremos ejemplos con más detalles y recomendaremos soluciones para algunos problemas comunes en el muestreo.

Archivos

Un gerente de reclutamiento y selección de una empresa quiere precisar si algunos datos que se dan en una solicitud de trabajo están correlacionados con el ausentismo del empleado. Es decir, si a partir de datos como edad, género, estado civil, nivel educativo y duración en otro trabajo, es factible predecir la conducta de ausentismo. Para establecer correlaciones se considerará como población a todas las personas contratadas durante 10 años. Se relacionan sus datos en la solicitud de empleo con los registros de faltas.

Como no hay una lista elaborada de estos individuos, el investigador decide acudir a los archivos de las solicitudes de empleo. Tales archivos constituyen su marco muestral a partir del cual se obtendrá la muestra. Calcula el tamaño de la población, obtiene el tamaño de la muestra y selecciona sistemáticamente cada elemento $1/K$, cada solicitud que será analizada. Aquí el problema que surge es que en el archivo hay solicitudes de gente que no fue contratada y, por tanto, no debe considerarse en el estudio.

En este caso, y en otros en los que no todos los elementos del marco de referencia o de una lista aparecen (por ejemplo, nombres en el directorio que no corresponden a una persona física), los especialistas en muestreo (Kish, 1995; Sudman, 1976) no aconsejan el reemplazo con el siguiente elemento, sino simplemente no tomar en cuenta ese elemento, es decir, hacer como si no existiera, y continuar con el intervalo de selección sistemática.

Mapas

Los mapas son muy útiles como marco de referencia en muestras de racimos. Por ejemplo, un investigador quiere saber qué motiva a los compradores de las tiendas de autoservicio. A partir de una lista de tiendas de cada cadena competidora marca sobre un mapa de la ciudad, todas las tiendas de autoservicios, las cuales constituyen una población de racimos, pues en cada tienda seleccionada entrevistará a un número de clientes. El mapa le permite ver la población (tiendas de autoservicio) y su situación geográfica, de manera que elige zonas donde coexistan diferentes tiendas competidoras, para asegurarse de que el consumidor de la zona tenga todas las posibles alternativas. En la actualidad hay mapas de todo tipo: mercadológicos, socioculturales, étnicos, marítimos, entre otros. El *Global Positioning System* (GPS) o Sistema de Posicionamiento Global ya puede ser muy útil para esta clase de muestreo.

Tamaño óptimo de una muestra

Tal como se mencionó, las muestras probabilísticas requieren dos procedimientos básicos: 1) la determinación del tamaño de la muestra y 2) la selección aleatoria de los elementos muestrales. Precisar adecuadamente el tamaño de la muestra puede tornarse complejo, esto depende del problema de investigación y la población a estudiar. Para el alumno y el lector en general, será muy útil comparar qué tamaño de muestra han empleado otros investigadores, a la luz de la revisión de la literatura. Para tal efecto, mostramos algunos ejemplos y reproducimos varias tablas (8.4, 8.5 y 8.6), que indican los tamaños de muestra más utilizados por los investigadores, según sus poblaciones (nacionales o regionales) y los subgrupos que quieren estudiarse en ellas.



Las muestras nacionales, es decir, las que representan a la población de un país, por lo común son de más de 1 000 sujetos. La muestra del estudio “¿Cómo somos los mexicanos?” (Hernández Medina, Narro y Rodríguez, 1987), consta de 1 737 sujetos repartidos de la siguiente manera:

Frontera y norte	696
Centro (sin la capital nacional o Distrito Federal)	426
Sur-sureste	316
Distrito Federal	299
	1 737

La muestra de los barómetros de opinión en España es nacional,⁸ incluye personas de ambos géneros, de 18 años o más y su tamaño es alrededor de 2 500 casos (Centro de Investigaciones Sociológicas, 2009). Su elección es por estratos y racimos. Primero, se eligen municipios y luego secciones. Los puntos de muestreo son 168 municipios y 49 provincias.

Los estratos están compuestos por siete categorías formadas por el cruce de las 17 comunidades autónomas con el tamaño del hábitat. Son las siguientes: *a*) menores o iguales a 2 000 habitantes, *b*) de 2 001 a 10 000, *c*) de 10 001 a 50 000, *d*) de 50 001 a 100 000, *e*) de 100 001 a 400 000, *f*) de 400 001 a un millón y *g*) más de un millón de habitantes (Berganza y García, 2005, p. 91).

En cambio, el Barómetro del Real Instituto Elcano (BRIE) en España comprende a 1 200 individuos (Real Instituto Elcano, 2009).

El Eurobarómetro es otra encuesta que abarca a diversos países de la Unión Europea (UE) y su muestra es de aproximadamente 1 000 personas por país, excepto en Alemania donde se consulta al doble y a Reino Unido, donde la *n* es igual a 1 300 (300 encuestas se efectúan en Irlanda) (Berganza y García, 2005). Por ejemplo, un Eurobarómetro sobre las mujeres y las elecciones europeas realizado el 4 de marzo del 2009, incluyó a 35 000 mujeres y 5 500 hombres de la UE (Oficina del Parlament Europeu a Barcelona, 2009).

En la tabla 8.4 observamos que el tipo de estudio en poco determina el tamaño de la muestra. Más bien, interviene en la decisión de que sean muestras nacionales o regionales.

Las muestras regionales (por ejemplo, las que representen al área metropolitana de la Ciudad de México u otra gran urbe con más de tres millones de habitantes), o de algún estado, departamento o provincia de un país, o algún municipio o región, son típicamente más pequeñas, con rangos de 400 a 700 individuos.

▲ **Tabla 8.4** Muestras utilizadas con frecuencia en investigaciones nacionales y regionales según área de estudio

Tipos de estudio	Nacionales	Regionales
Económicos	1 000+	100
Médicos	1 000+	500
Conductas	1 000+	700-300
Actitudes	1 000+	700-400
Experimentos de laboratorio	---	100

⁸ Los barómetros son *surveys* o encuestas de alcance nacional o continental e incluyen cuestiones políticas, económicas, sociales y de actualidad.

El tamaño de una muestra depende también del número de subgrupos que nos interesan en una población. Por ejemplo, podemos subdividirla en hombres y mujeres de cuatro grupos de edad o, aún más, en hombres y mujeres de cuatro grupos de edad en cada uno de cinco niveles socioeconómicos. Si éste fuera el caso estaríamos hablando de 40 subgrupos y, por ende, de una muestra mayor. En la tabla 8.5 se describen muestras típicas de acuerdo con los subgrupos bajo estudio, según su cobertura (estudios nacionales o estudios especiales o regionales) y según su unidad de análisis; es decir, se trata de individuos o de organizaciones. En esta última instancia el número de la muestra se reduce, ya que casi siempre representa una gran fracción de la población total.

▲ **Tabla 8.5** Muestras típicas de estudios sobre poblaciones humanas y organizaciones

Número de subgrupos	Población de individuos u hogares		Población de organizaciones	
	Nacionales	Regionales	Nacionales	Regionales
Ninguno-pocos (menos de 5)	1 000-1 500	200-500	200-500	50-200
Promedio (5 a 10)	1 500-2 500	500-1 000	500-1 000	200-500
Más de 10	2 500 +	1 000 +	1 000 +	500 +

Otra tabla que nos ayuda a comprender el tema que estamos analizando es la 8.6, la cual se basa en Mertens (2005, p. 327) y Borg y Gall (1989), de acuerdo con el propósito del estudio. Aquí cada número es el mínimo sugerido.

▲ **Tabla 8.6** Tamaños de muestra mínimos en estudios cuantitativos

Tipo de estudio	Tamaño mínimo de muestra
Transeccional descriptivo o correlacional	30 casos por grupo o segmento del universo.
Encuesta a gran escala	100 casos para el grupo o segmento más importante del universo y de 20 a 50 casos para grupos menos importantes.
Causal	15 casos por variable independiente.
Experimental o cuasiexperimental	15 por grupo.

Las tablas 8.4 a 8.6 se construyeron con base en artículos de investigación publicados en revistas especializadas y en Sudman (1976), y nos dan una idea de las muestras que utilizan otros investigadores, de manera que le ayudarán a establecer el tamaño de su muestra. En el caso de los experimentos, la muestra representa el balance entre un mayor número de casos y el número que podamos manejar. Recordemos que algunas pruebas estadísticas exigen 15 casos como mínimo por grupo de comparación (Mertens, 2005).

Repasemos que lo óptimo de una muestra depende de cuánto se aproxima su distribución a la distribución de las características de la población. Esta aproximación mejora al incrementarse el tamaño de la muestra. La “normalidad” de la distribución en muestras grandes no obedece a la normalidad de la distribución de una población. La distribución de diversas variables a veces es “normal” y en ocasiones está lejos de serlo. Sin embargo, la distribución de muestras de 100 o más elementos tiende a ser normal y esto sirve para el propósito de hacer estadística inferencial (generalizar de la muestra al universo). A lo anterior se le llama **teorema del límite central**.

Distribución normal: esta distribución en forma de campana se logra generalmente con muestras de 100 o más unidades muestrales, y es útil y necesaria cuando se hacen inferencias de tipo estadístico.

Teorema del límite central Señala que una muestra de más de cien casos será una muestra con una distribución normal en sus características, lo cual sirve para el propósito de hacer estadística inferencial.

¿Cómo y cuáles son las muestras no probabilísticas?

Las muestras no probabilísticas, también llamadas muestras dirigidas, suponen un procedimiento de selección informal. Se utilizan en diversas investigaciones cuantitativas y cualitativas. No las revisaremos

ahora, sino en el capítulo 13 “Muestreo cualitativo”. Por el momento comentaremos que seleccionan individuos o casos “típicos” sin intentar que sean representativos de una población determinada. Por ello, para fines deductivos-cuantitativos, donde la generalización o extrapolación de resultados hacia la población es una finalidad en sí misma, las muestras dirigidas implican algunas desventajas. La primera es que, al no ser probabilísticas, no es posible calcular con precisión el error estándar, es decir, no podemos calcular con qué nivel de confianza hacemos una estimación. Esto es un grave inconveniente si consideramos que la estadística inferencial se basa en la teoría de la probabilidad, por lo que las pruebas estadísticas en muestras no probabilísticas tienen un valor limitado a la muestra en sí, mas no a la población. Es decir, los datos no pueden generalizarse a ésta. En las muestras de este tipo, la elección de los casos no depende de que todos tengan la misma probabilidad de ser elegidos, sino de la decisión de un investigador o grupo de personas que recolectan los datos.

La única ventaja de una muestra no probabilística —desde la visión cuantitativa— es su utilidad para determinado diseño de estudio que requiere no tanto una “representatividad” de elementos de una población, sino una cuidadosa y controlada elección de casos con ciertas características especificadas previamente en el planteamiento del problema.

Para el enfoque cualitativo, al no interesar tanto la posibilidad de generalizar los resultados, las muestras no probabilísticas o dirigidas son de gran valor, pues logran obtener los casos (personas, contextos, situaciones) que interesan al investigador y que llegan a ofrecer una gran riqueza para la recolección y el análisis de los datos.

Muestreo al azar por marcado telefónico (*Random Digit Dialing*)

Ésta es una técnica que los investigadores utilizan para seleccionar muestras telefónicas. Involucra identificar áreas geográficas —para ser muestreadas al azar— y sus correspondientes códigos telefónicos e intercambios (los primeros dígitos del número telefónico que las identifican). Luego, los demás dígitos del número a marcar pueden ser generados al azar de acuerdo con los casos que requerimos para la muestra (n). Es posible reconocer qué intercambios son usados de forma primaria para teléfonos residenciales y enfocar el muestreo en ese subgrupo. Asimismo, es muy útil para incluir en muestras a teléfonos celulares o móviles (Hernández Sampieri y Mendoza, 2008).

Para mayores referencias de esta técnica recomendamos Fowler (2002) y Link, Town y Mokdad (2007). Un excelente ejemplo para ver cómo se conforma una muestra mediante este método se puede encontrar en Williams, Van Dyke y O’Leary, (2006).

Una máxima del muestreo y el alcance del estudio

Ya sea que se trate de un tipo de muestreo u otro, lo importante es elegir a los informantes (o casos) adecuados, de acuerdo con el planteamiento del problema y lograr el acceso a ellos.

Los estudios exploratorios regularmente emplean muestras dirigidas, aunque podrían usarse muestras probabilísticas. Las investigaciones experimentales, la mayoría de las veces utilizan muestras dirigidas, porque como se comentó, es difícil manejar grupos grandes (debido a ello se ha insistido que, en los experimentos, la validez externa se consolida mediante la repetición o reproducción del estudio). Los estudios no experimentales descriptivos o correlacionales-causales deben emplear muestras probabilísticas si quieren que sus resultados sean generalizados a una población.

Asimismo, en ocasiones la muestra puede ser en varias etapas (polietápica). Por ejemplo, primero elegir universidades, luego, escuelas o facultades, después, salones o grupos y finalmente, estudiantes.



Resumen



- En el capítulo se definió el concepto de muestra.
- Además, se describió cómo seleccionar una muestra en el proceso cuantitativo. Lo primero que se debe plantear es sobre qué o quiénes se van a recolectar los datos, lo cual corresponde a precisar la unidad de análisis. Después, se procede a delimitar claramente la población, con base en los objetivos del estudio y en cuanto a características de contenido, de lugar y de tiempo.
- La muestra es un subgrupo de la población y puede ser probabilística o no probabilística.
- Elegir qué tipo de muestra se requiere depende del enfoque y alcances de la investigación, los objetivos del estudio y el diseño.
- En el enfoque cuantitativo las muestras probabilísticas son esenciales en diseños de investigación por encuestas, donde se pretenden generalizar los resultados a una población. La característica de este tipo de muestras es que todos los elementos de la población al inicio tienen la misma probabilidad de ser elegidos. Así, los elementos muestrales tendrán valores muy aproximados a los valores de la población, ya que las mediciones del subconjunto serán estimaciones muy precisas del conjunto mayor. Tal precisión depende del error de muestreo, llamado también error estándar.
- Para una muestra probabilística necesitamos dos elementos: determinar el tamaño adecuado de la muestra y seleccionar los elementos muestrales en forma aleatoria.
- El tamaño de la muestra se calcula mediante fórmulas o por medio del programa STATS®, que se encuentra en el CD que acompaña al libro.
- Las muestras probabilísticas son: simples, estratificadas, sistemáticas y por racimos. La estratificación aumenta la precisión de la muestra e implica el uso deliberado de submuestras para



cada estrato o categoría que sea relevante en la población. Muestrear por racimos o conglomerados implica diferencias entre la unidad de análisis y la unidad muestral. En este tipo de muestreo hay una selección en varias etapas, todas con procedimientos probabilísticos. En la primera se seleccionan los racimos y dentro de los racimos, a los participantes que van a ser medidos.

- Los elementos muestrales de una muestra probabilística siempre se eligen aleatoriamente para asegurarnos de que cada elemento tenga la misma probabilidad de ser seleccionado. Es posible utilizar cuatro procedimientos de selección: 1) tómbola, 2) números aleatorios, 3) uso del subprograma de números aleatorios del STATS® y 4) selección sistemática. Todo procedimiento de selección depende de listados o bases de datos, ya sea existentes o construidas ad hoc. Los listados pueden ser: la guía telefónica, listas de asociaciones, listas de escuelas oficiales, etc. Cuando no existen listas de elementos de la población, se recurre a otros marcos de referencia que contengan descripciones del material, organizaciones o participantes seleccionados como unidades de análisis. Algunos de éstos pueden ser archivos, hemerotecas y mapas, así como internet.
- Las muestras no probabilísticas pueden también llamarse muestras dirigidas, pues la elección de casos depende del criterio del investigador.
- En el teorema del límite central se señala que una muestra de más de cien casos será una muestra con una distribución normal en sus características; sin embargo, la normalidad no debe confundirse con probabilidad. Mientras lo primero es necesario para efectuar pruebas estadísticas, lo segundo es requisito indispensable para hacer inferencias correctas sobre una población.



Conceptos básicos

Base de datos
Elementos muestrales
Error estándar
Muestra
Muestra no probabilística o dirigida
Muestra probabilística
Nivel deseado de confianza
Población

Representatividad
Selección aleatoria
Selección sistemática
Tamaño de muestra
Teorema del límite central
Unidad de análisis
Unidad muestral



Ejercicios

- Se forman grupos de tres o cuatro personas. Cada grupo dispone de 15 minutos para formular una pregunta de investigación. El problema puede ser de cualquier área de estudio. Lo que conviene aquí es que sea sobre un tema que realmente inquiete a los estudiantes, algo que ellos consideren un fenómeno importante. Las preguntas de investigación se van anotando en el pizarrón. Después y junto a cada una de estas preguntas se define quiénes van a ser medidos. Discutir por qué sí y por qué no son correctas las respuestas de los estudiantes.
- Como secuencia del ejercicio anterior se proponen los siguientes temas de investigación. Supongamos que, en otro curso, estudiantes de un taller de investigación sugirieron los siguientes temas para investigar. En cada caso señalar quiénes van a ser medidos, para lograr resultados en las investigaciones propuestas.
 - Tema 1. ¿Qué efecto tienen los anuncios de bebidas alcohólicas sobre los jóvenes?
 - Tema 2. Hace tres meses que se implantó en una fábrica de motores un programa de círculos de calidad. ¿Ha tenido éxito dicho programa?
 - Tema 3. ¿Los niños que cursaron la primaria en escuelas laicas y mixtas tienen un mejor desempeño académico en la universidad que los que provienen de escuelas religiosas de un solo género?
 - Tema 4. ¿Qué diferencias existen entre los comerciales de champú de la televisión española, la argentina y la venezolana?
- Seleccione dos estudios de alguna publicación científica (vea en el CD anexo: Material complementario → Apéndices → Apéndice 1) y/o dos tesis. Analice los siguientes aspectos:
 - ¿Cuál es el problema de investigación?
 - ¿Cuál es la muestra?
 - ¿Cómo fue elegida?
 - ¿Son adecuadas la muestra y el procedimiento de muestreo para el problema que se investigó?
 - ¿Cuáles son los principales resultados o conclusiones?
 - ¿Dichos resultados son generalizables a una población mayor?
 - Con base en la muestra, ¿pueden tomarse como serias dichas generalizaciones? Evalúe la solidez de los cuatro estudios, tomando como criterios los aspectos *a*, *b*, *c*, *d*, *e*, *f* y *g*.
- Supongamos que trabaja en un despacho que realiza investigaciones y que diversos clientes le solicitan que los asesore en estudios de diferente índole. ¿Qué tipo de muestra sugeriría para cada uno? Fundamente su sugerencia.

Cliente	Necesidad	Tipo de muestra
4.1 Clínica de terapias psicoemocionales	Pacientes con cáncer que siguen la terapia reaccionan mejor a los tratamientos médicos usuales que los enfermos de cáncer que no toman la terapia.	
4.2 Empresa en el giro químico	Definir cuáles son nuestros empleados y obreros, anteriores y presentes, que tienen menos ausentismo, es decir, ¿hay un perfil del ausentismo?	
4.3 Empresa de cosmología	¿Qué nociones tienen las jóvenes (de 15 a 20 años) sobre su arreglo personal y el cuidado de su cutis? ¿Funcionaría crear una línea de productos exclusivamente para ellas?	
4.4 Grupo que defiende los derechos del consumidor	¿Qué quejas tienen los niños sobre los juguetes del mercado?, ¿se rompen?, ¿son peligrosos?, ¿aburridos?, ¿cuál es su durabilidad?, etcétera.	
4.5 Partidos políticos	¿Por cuál candidato a gobernador votarán los ciudadanos de determinado estado o provincia?	

- Supongamos que una asociación iberoamericana de profesionales cuenta con 5 000 miembros. La junta directiva ha decidido hacer una encuesta (por teléfono o por correo electrónico) a los asociados para indagar, entre otras cosas, lugar de trabajo, puesto que ocupan, salario aproximado, licenciatura cursada, generación, estudios posteriores, oportunidades de avance percibidas, etc. En resumen, se piensa publicar un perfil profesional actualizado con el propósito de retroalimentar a los asociados. Como



sería muy costoso llegar a los 5000 miembros repartidos en España, América Latina y Estados Unidos, ¿qué tamaño de muestra se necesita si queremos un error estándar no mayor de 0.015?

Una vez definido el tamaño de la muestra, ¿cómo sería el proceso de selección a fin de que los resultados obtenidos con base en la muestra sean generalizables a toda la población? Es decir, se pretende reportar un perfil certero de los 5000 socios de dicha asociación profesional.

6. Seleccione un tamaño de muestra adecuado para su institución, mediante el STATS®.

7. Con respecto al ejemplo de estudio que ha venido desarrollando en el proceso cuantitativo, piense cómo seleccionaría la muestra apropiada de acuerdo con su planteamiento, objetivos, hipótesis y diseño. ¿Cuál sería el universo o población, la unidad de análisis y el procedimiento de selección? y ¿qué tamaño tendría la muestra?

Recuerde ver las respuestas a los ejercicios en el CD anexo: Material complementario → Apéndices → Apéndice 3.



Ejemplos desarrollados

La televisión y el niño

Para el estudio, primero se realizó un análisis exploratorio y una prueba piloto con 60 niños de diversos estratos socioeconómicos. Con base en ello se corrigió el cuestionario para proceder al estudio definitivo.

1. Límites de población:

Todos los niños del área metropolitana de la Ciudad de México, que cursen 4o., 5o. y 6o. de primaria en escuelas privadas y públicas del turno matutino.

2. Proceso de selección:

Se estableció una muestra probabilística estratificada por racimo, donde en una primera etapa se seleccionaron escuelas para, en última instancia, llegar a los niños. La muestra se obtuvo de una base de datos de la Secretaría de Educación Pública, que contuviera listadas e identificadas a todas las escuelas primarias del área metropolitana de la ciudad de México.

Se excluyó a escuelas del turno vespertino y las diseñadas para niños con capacidades diferentes o habilidades especiales. La selección también estratificó el nivel socioeconómico en cuatro categorías: A, B, C y D (de acuerdo con los criterios del mapa mercadológico de la ciudad de México, A = ingresos familiares elevados, B = medios, C = medios bajos y D = bajos). Por tanto, se eligieron las escuelas de los siguientes estratos:

1. escuelas públicas clase A;
2. escuelas privadas clase A;
3. escuelas públicas clase B;
4. escuelas privadas clase B;
5. escuelas públicas clase C;
6. escuelas privadas clase C;

7. escuelas públicas clase D;
8. escuelas privadas clase D.

Cada lista representó un estrato de la población y de cada una de ellas se seleccionó una muestra de escuelas: A, B, C, D, que representan niveles socioeconómicos. Posteriormente, de cada escuela se eligieron los niños para conformar la muestra final.

Una vez hechos los cálculos, se determinó que de cada estrato se seleccionaran cuatro escuelas, es decir n es igual a 32 escuelas ubicadas en diversas colonias que incluyeron a todas las delegaciones (municipalidades). En la segunda etapa se seleccionaron por muestreo aleatorio simple los niños de cada escuela. En el ejemplo, 264 infantes por escuela de 4o., 5o. y 6o. grados (88 por cada uno). Una muestra total de 2112 que implicó ajustes y reemplazos.

La pareja y relación ideales

Para conocer el tamaño del universo, se obtuvo información proporcionada por la Asociación Nacional de Universidades e Instituciones de Educación Superior, Federación de Instituciones Mexicanas Particulares de Educación Superior y el gobierno de Guanajuato. Asimismo, se acudió a fuentes electrónicas (páginas web de las instituciones) y se solicitó directamente el dato a las organizaciones educativas involucradas. El tamaño de la población total es de aproximadamente 13000 estudiantes.⁹ Utilizando

⁹ No se proporciona la matrícula de cada institución en particular, porque cuatro universidades solicitaron expresamente que no se difundiera el dato. También, cabe mencionar que el tamaño del universo es aproximado debido a que hasta el final del semestre se tiene información precisa de las bajas escolares.

el STATS®, tendríamos que un tamaño de muestra adecuado para esta población (95% de confianza, 5% de error y $p = 0.5$ o 50%) es de 373 casos. Sin embargo, se prefirió segmentar al universo en: 1) instituciones con matrícula considerable (más de 2000 alumnos) y 2) universidades con matrícula estándar para una ciudad intermedia (1000 a 1500 estudiantes). En el primer estrato estuvieron dos organizaciones (que representa un total de 6000 universitarios) y en el segundo siete (7000 alumnos). Cada estrato fue concebido como una población y entonces se calculó el tamaño de muestra mediante STATS®, el resultado fue: estrato 1 ($n = 361$), estrato 2 ($n = 364$). Así, para el estrato 1 se consideró entrevistar en una institución a 180 universitarios y en la otra a 181. En el caso del estrato 2, se administró el instrumento de medición en cada una de las siete universidades a 52 estudiantes. En un futuro se agregará al estudio al Instituto Tecnológico Roque, al Centro Universitario ITESBA y a otras organizaciones, para poder comparar entre instituciones y cada una podría concebirse como una población en sí misma.

El abuso sexual infantil

El abuso sexual infantil

El estudio es un experimento y la muestra es dirigida. Se reclutaron preescolares de tres centros de desarrollo infantil con una población similar, hijas e hijos de madres que laboran para la Secretaría de Educación Pública del Estado de Querétaro. Se evaluaron seis grupos escolares que fueron asignados a tres grupos experimentales ($n_1 = 49$ niños, $n_2 = 22$ niños y $n_3 = 79$ niños).

Al inicio del proceso se obtuvo anuencia de las autoridades escolares de los centros. En general, se hicieron reuniones previas con los padres de familia para informarles del programa. Se efectuó una sesión de acercamiento en la cual, la persona que aplicó las escalas se presentó con los niños y las niñas, asimismo, desarrolló actividades lúdicas para establecer confianza y cercanía con los grupos, además les explicó de forma general el proceso a llevarse a cabo y su participación fue de carácter voluntario (tenían la posibilidad de negarse). Antes de cada evaluación, se les pidió su consentimiento a todos los infantes.



Los investigadores opinan

La importancia de la investigación radica en que genera conocimientos, lo cual contribuye al desarrollo social. Por consiguiente, es importante que los estudiantes tengan el gusto e interés profesional por investigar.

A partir de la preferencia por determinado tema, se desprende la orientación que se le debe dar al proyecto, donde tiene que haber claridad conceptual y

exactitud en la aproximación al problema, además de procurar la comunicación de los resultados.

ÁLVARO CAMACHO MEDINA

Docente

FACULTAD DE MERCADERO Y PUBLICIDAD

Politécnico Grancolombiano

Bogotá, Colombia

En nuestra realidad existen investigaciones serias que aportan indicadores de cómo se encuentran, por ejemplo, los diferentes niveles del sistema educativo peruano; sin embargo, no son suficientes en la aplicación de propuestas metodológicas, ya sea por la selección de la muestra, el empleo de instrumentos adecuados o la preparación del personal que las lleva a cabo.

Por tal razón, quienes tenemos la responsabilidad de orientar proyectos debemos infundir a nuestros alumnos que la investigación es un proceso que convoca nuestra energía y perseverancia para obtener

resultados que sean significativos para la sociedad peruana.

Para ello se requiere vivir determinadas experiencias. En el caso de la educación, sería recomendable visitar un centro académico que ensaye diferentes y nuevos enfoques para conocer el medio, dialogar con los protagonistas y descubrir su problemática.

Un buen planteamiento del problema nos permitirá orientar la investigación, precisar las variables a analizar, conocer el grupo con el que se pretende trabajar, determinar los objetivos y, en un momento dado, redactar coherentemente los resultados.

Por último, considero que los resultados de una investigación se tornan significativos cuando, además de presentar datos cuantitativos, en ella se consideran también datos cualitativos. Una experiencia de investigación debe tomar en cuenta ambos enfoques, porque así será posible admirarla y apreciarla de forma integral.

ING. GUILLERMO EVANGELISTA BENITES

Docente principal

FACULTAD DE INGENIERÍA QUÍMICA

Universidad Nacional de Trujillo

Trujillo, Perú

**Proceso de investigación cuantitativa****Paso 8 Recolectar los datos**

- Definir la forma idónea de recolectar los datos de acuerdo con el planteamiento del problema y las etapas previas de la investigación.
- Seleccionar o elaborar uno o varios instrumentos o métodos para recolectar los datos requeridos.
- Aplicar los instrumentos o métodos.
- Obtener los datos.
- Codificar los datos.
- Archivar los datos y prepararlos para su análisis por computadora.

Objetivos del aprendizaje

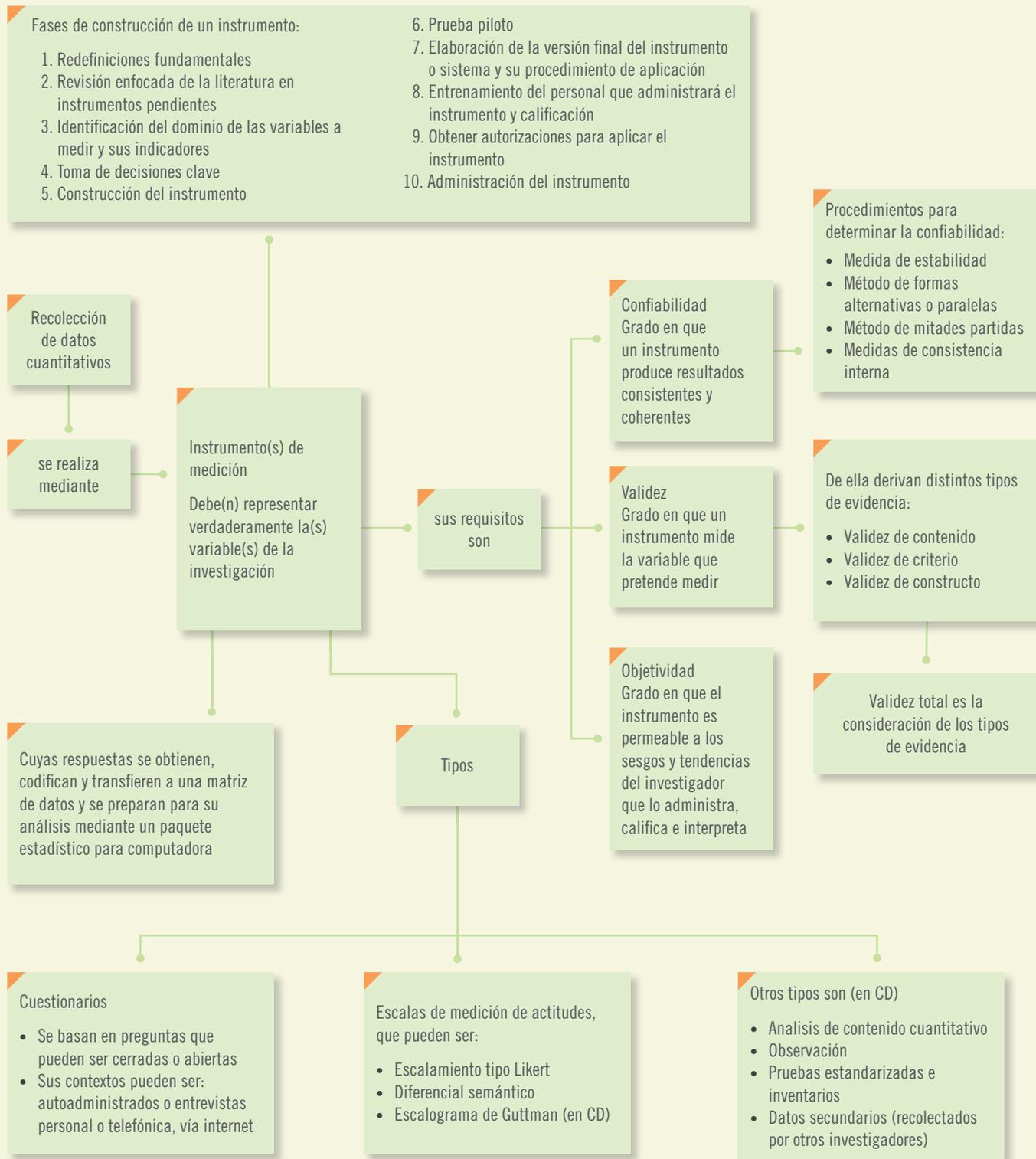
Al terminar este capítulo, el alumno será capaz de:

- 1 Visualizar diferentes métodos para recolectar datos cuantitativos.
- 2 Entender el significado de “medir” y su importancia en el proceso cuantitativo.
- 3 Comprender los requisitos que toda recolección de datos debe incluir.
- 4 Conocer los principales instrumentos para recolectar datos cuantitativos.
- 5 Elaborar y aplicar los diferentes instrumentos de recolección de datos cuantitativos.
- 6 Preparar los datos para su análisis cuantitativo.

Síntesis

En el capítulo se analizan los requisitos que un instrumento debe cubrir para recolectar apropiadamente datos cuantitativos: confiabilidad, validez y objetividad. Asimismo, se define el concepto de medición y los errores que pueden cometerse al recolectar datos.

A lo largo del capítulo se presenta el proceso para elaborar un instrumento de medición y las principales alternativas para recolectar datos: cuestionarios y escalas de actitudes. Por último, se examina el procedimiento de codificación de datos cuantitativos y la forma de prepararlos para su análisis.



Nota: El capítulo se termina de integrar con otro del CD anexo (Material complementario → Capítulos → Capítulo 7), titulado “Recolección de los datos cuantitativos: segunda parte”, que contiene otras alternativas de instrumentos para recolectar datos como el análisis de contenido y los sistemas de observación (en ediciones anteriores se localizaba en este mismo capítulo), además de pruebas e inventarios, escalograma de Guttman (escala de actitudes) y datos secundarios.



¿Qué implica la etapa de recolección de datos?

OQ1 Una vez que seleccionamos el diseño de investigación apropiado y la muestra adecuada (probabilística o no probabilística), de acuerdo con nuestro problema de estudio e hipótesis (si es que se establecieron), la siguiente etapa consiste en recolectar los datos pertinentes sobre los atributos, conceptos o variables de las unidades de análisis o casos (participantes, grupos, organizaciones, etcétera).

Recolectar los datos implica elaborar un plan detallado de procedimientos que nos conduzcan a reunir datos con un propósito específico. Este plan incluye determinar:

- a) ¿Cuáles son las fuentes de donde se obtendrán los datos? Es decir, los datos van a ser proporcionados por personas, se producirán de observaciones o se encuentran en documentos, archivos, bases de datos, etcétera.
- b) ¿En dónde se localizan tales fuentes? Regularmente en la muestra seleccionada, pero es indispensable definir con precisión.
- c) ¿A través de qué medio o método vamos a recolectar los datos? Esta fase implica elegir uno o varios medios y definir los procedimientos que utilizaremos en la recolección de los datos. El método o métodos deben ser confiables, válidos y objetivos.
- d) Una vez recolectados, ¿de qué forma vamos a prepararlos para que puedan analizarse y respondamos al planteamiento del problema?

El plan se nutre de diversos elementos:

1. Las *variables*, conceptos o atributos a medir (contenidos en el planteamiento e hipótesis o directrices del estudio).
2. Las *definiciones operacionales*. La manera como hemos operacionalizado las variables es crucial para determinar el método para medirlas, lo cual a su vez, resulta fundamental para realizar las inferencias de los datos.
3. La *muestra*.
4. Los *recursos disponibles* (de tiempo, apoyo institucional, económicos, etcétera).

Desde luego, aquí hemos simplificado la información por motivos de espacio.

El plan se implementa para obtener los datos requeridos, no olvidemos que todos los atributos, cualidades y variables deben ser medibles. Un ejemplo de plan de este tipo se puede ver en la figura 9.1.

Con la finalidad de recolectar datos disponemos de una gran variedad de instrumentos o técnicas, tanto cuantitativas como cualitativas, es por ello que en un mismo estudio podemos utilizar ambos tipos. Incluso, hay instrumentos como la prueba de propósito vital (PIL) (que evalúa el propósito de vida de una persona) de Crumbaugh y Maholick (1969) que contienen una parte cuantitativa y una cualitativa (Brown, Ashcroft y Miller, 1998). Esto se revisará en el capítulo 17: “Los métodos mixtos”.

Antes de continuar es necesario revisar algunos conceptos esenciales para la recolección de los datos cuantitativos.

¿Qué significa medir?

OQ2 En la vida diaria medimos constantemente. Por ejemplo, al levantarnos por las mañanas, miramos el reloj despertador y “medimos” la hora; al bañarnos, *ajustamos* la temperatura del agua en la tina o la regadera, *calculamos* la cantidad de café que habremos de colocar en la cafetera; nos asomamos por la ventana y estimamos cómo será el día para decidir la ropa o atuendos que nos pondremos; al ver el tránsito desde el autobús u otro vehículo, evaluamos e inferimos a qué hora llegaremos a la universidad o al trabajo, así como la velocidad a la que transitamos (u observamos el velocímetro); en ocasiones contamos cuántos anuncios espectaculares observamos en el trayecto u otras cuestiones, incluso inferimos, a partir de ciertos signos, acerca del operador del autobús u otros conductores: ¿qué tan alegres o enojados están?, además de otras actividades. Medir es parte de nuestras vidas (Bostwick y Kyte, 2005).

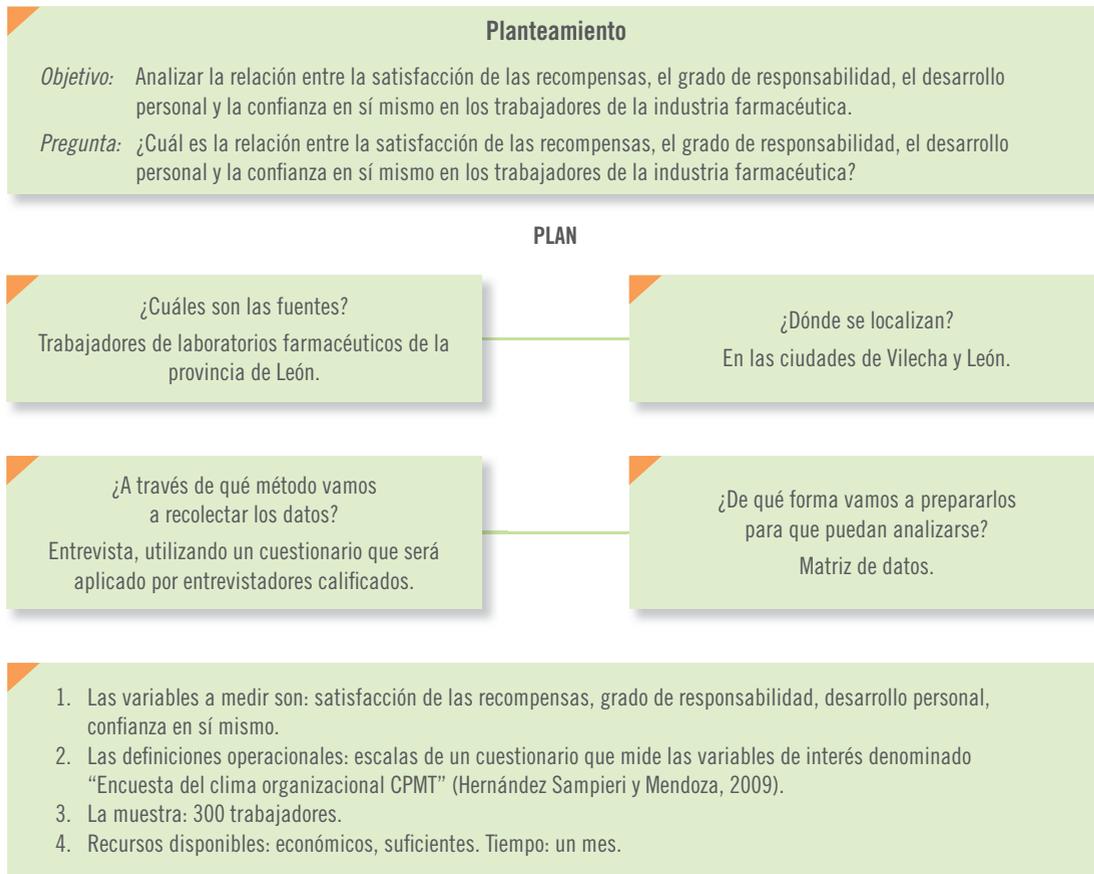


Figura 9.1 Ejemplo de plan para la obtención de datos.

De acuerdo con la definición clásica del término, ampliamente difundida, **medir** significa “asignar números, símbolos o valores a las propiedades de objetos o eventos de acuerdo con reglas” (Stevens, 1951). Desde luego, no se asignan a los objetos, sino a sus propiedades (Bostwick y Kyte, 2005). Sin embargo, como señalan Carmines y Zeller (1991), 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, ya que son demasiado abstractos para ello. La disonancia cognitiva, la pareja ideal, el clima organizacional, la cultura fiscal y la credibilidad son conceptos tan abstractos que no deben ser considerados “cosas que pueden verse o tocarse” (definición de objeto) o solamente como “resultado, consecuencia o producto” (definición de evento) (Carmines y Zeller, 1991). Este razonamiento nos hace sugerir que es más adecuado definir la **medición** como “el proceso de vincular conceptos abstractos con indicadores empíricos”, el cual se realiza mediante un plan explícito y organizado para clasificar (y con frecuencia cuantificar) los datos disponibles (los indicadores), en términos del concepto que el investigador tiene en mente (Carmines y Zeller, 1991). En este proceso, el instrumento de medición o de recolección de datos tiene un papel central. Sin él, no hay observaciones clasificadas.

Medición Proceso que vincula conceptos abstractos con indicadores empíricos.

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 se representa por medio de la respuesta. Así, los registros del instrumento de medición representan valores visibles de conceptos abstractos. Un **instrumento de**

Instrumento de medición Recurso que utiliza el investigador para registrar información o datos sobre las variables que tiene en mente.

medición adecuado es aquel que registra datos observables que representan verdaderamente los conceptos o las variables que el investigador tiene en mente (Grinnell, Williams y Unrau, 2009). En términos cuantitativos: capturo verdaderamente la “realidad” que deseo capturar. Bostwick y Kyte (2005) lo señalan de la siguiente forma: “La función de la medición es establecer una correspondencia entre el “mundo real” y el “mundo conceptual”. El primero provee evidencia empírica, el segundo proporciona modelos teóricos para encontrar sentido a ese segmento del mundo real que estamos tratando de describir.

En toda investigación cuantitativa 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 datos en realidad representa a las variables que tenemos en mente. Si no es así, nuestra medición es deficiente; por tanto, la investigación no es digna de tomarse en cuenta. Desde luego, no hay medición perfecta. Es casi imposible que representemos con fidelidad variables tales como la inteligencia emocional, la motivación, el nivel socioeconómico, el liderazgo democrático, el abuso sexual infantil 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. Se trata de un precepto básico del enfoque cuantitativo. Al medir estandarizamos y cuantificamos los datos (Bostwick y Kyte, 2005; Babbie, 2009).

¿Qué requisitos debe cubrir un instrumento de medición?

QA3 Toda medición o instrumento de recolección de datos debe reunir tres requisitos esenciales: *confiabilidad, validez y objetividad*.

La confiabilidad

Confiabilidad Grado en que un instrumento produce resultados consistentes y coherentes.

La **confiabilidad** de un instrumento de medición se refiere al grado en que su aplicación repetida al mismo individuo u objeto produce resultados iguales. Por ejemplo, si se midiera en este momento la temperatura ambiental usando un termómetro y éste indicara que hay 22°C, y un minuto más tarde se consultara otra vez y señalara 5°C, tres minutos después se observara nuevamente y éste indicara 40°C, dicho termómetro no sería confiable, ya que su aplicación repetida produce resultados distintos. Asimismo, si una prueba de inteligencia (Intelligence Quotient, IQ) se aplica hoy a un grupo de personas y da ciertos valores de inteligencia, se aplica un mes después y proporciona valores diferentes, al igual que en subsecuentes mediciones, tal prueba no sería confiable (analice los valores de la tabla 9.1, suponiendo que los coeficientes de inteligencia oscilaran entre 100 y 135). Los resultados no son coherentes, pues no se puede “confiar” en ellos.

▲ **Tabla 9.1** Ejemplo de resultados proporcionados por un instrumento de medición sin confiabilidad

Primera aplicación		Segunda aplicación		Tercera aplicación	
Mariana	135	Sergio	131	Guadalupe	127
Viridiana	125	Laura	130	Agustín	120
Sergio	118	Chester	125	Mariana	118
Laura	110	Guadalupe	112	Laura	115
Guadalupe	108	Mariana	110	Chester	112
Chester	106	Viridiana	105	Viridiana	108
Agustín	100	Agustín	101	Sergio	105

La confiabilidad de un instrumento de medición se determina mediante diversas técnicas, las cuales se comentarán brevemente después de revisar los conceptos de validez y objetividad.

La 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 válido para medir la inteligencia debe medir la inteligencia y no la memoria. Un método para medir el rendimiento bursátil tiene que medir precisamente esto y no la imagen de una empresa. En apariencia 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 género al que pertenece una persona). Sin embargo, la situación no es tan simple cuando se trata de variables como la motivación, la calidad del servicio a los clientes, la actitud hacia un candidato político, y menos aún con sentimientos y emociones, así como de otras variables con las que trabajamos en todas las ciencias. 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 de la validez: ¿está midiendo lo que cree que está midiendo? Si es así, su medida es válida; si no, evidentemente carece de validez.

La validez es un concepto del cual pueden tenerse diferentes tipos de evidencia (Gronlund, 1990; Streiner y Norman, 2008; Wiersma y Jurs, 2008; y Babbie, 2009): 1) *evidencia relacionada con el contenido*, 2) *evidencia relacionada con el criterio* y 3) *evidencia relacionada con el constructo*. A continuación analizaremos 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 el que la medición representa al concepto o variable medida (Bohrstedt, 1976). Por ejemplo, una prueba de operaciones aritméticas no tendrá validez de contenido si incluyera sólo problemas de resta y excluyera problemas de suma, multiplicación o división. O bien, una prueba de conocimientos sobre las canciones de Los Beatles no deberá basarse solamente en sus álbumes *Let it Be* y *Abbey Road*, sino que debe incluir canciones de todos sus discos. O una prueba de conocimientos de líderes históricos de América Latina que omita a Simón Bolívar, Salvador Allende o Benito Juárez, y se concentre en Eva y Domingo Perón, Augusto Pinochet, el cura Miguel Hidalgo y otros líderes.

Un instrumento de medición requiere tener representados prácticamente a todos o la mayoría de los componentes del dominio de contenido de las variables a medir. Este hecho se ilustra en la figura 9.2.

Validez Grado en que un instrumento en verdad mide la variable que se busca medir.

Validez de contenido Se refiere al grado en que un instrumento refleja un dominio específico de contenido de lo que se mide.



Figura 9.2 Ejemplo de un instrumento de medición con validez de contenido frente a otro que carece de ella.

El dominio de contenido de una variable normalmente está definido o establecido por la literatura (teoría y estudios antecedentes). En estudios exploratorios donde las fuentes previas son escasas, el investigador comienza a adentrarse en el problema de investigación y a sugerir cómo puede estar cons-

titudino tal dominio. De cualquier manera en cada estudio uno debe probar que el instrumento utilizado es válido. Un ejemplo del intento por establecer el dominio de contenido de una variable es el siguiente:

EJEMPLO

Hernández Sampieri (2005), para establecer el dominio de la variable clima organizacional, revisó 20 estudios clásicos sobre el concepto, comprendidos entre 1964 y 1977, así como más de 100 investigaciones publicadas en revistas científicas entre 1975 y 2005. Por otro lado, consideró diversos libros sobre el tema, tres metaanálisis y otras tantas revisiones del estado del conocimiento sobre dicho clima. También evaluó 15 estudios efectuados en el contexto donde habría de llevar a cabo su propia investigación. Encontró que en la literatura se han considerado decenas de dimensiones o componentes del clima organizacional, por lo que realizó un análisis para determinar cuáles habían sido los más frecuentes, éstos fueron: 1) moral, 2) apoyo de la dirección, 3) innovación, 4) identificación con la empresa, 5) comunicación, 6) percepción del desempeño, 7) motivación intrínseca, 8) autonomía, 9) satisfacción general, 10) liderazgo, 11) visión y 12) recompensas o retribución. Dejó a un lado otras, como confianza en sí mismo, estándares de excelencia o conformidad. De lo anterior generó su instrumento de medición.

Si el dominio de un instrumento es demasiado estrecho con respecto al dominio de la variable, el primero no representará a ésta. La pregunta que se responde con la validez de contenido es: *¿el instrumento mide adecuadamente las principales dimensiones de la variable en cuestión?* En un cuestionario, por ejemplo, cabría interrogar: *¿las preguntas qué tan bien representan a todas las preguntas que pudieran hacerse?*

2. Evidencia relacionada con el criterio

Validez de criterio Se establece al validar un instrumento de medición al compararlo con algún criterio externo que pretende medir lo mismo.

La **validez de criterio** establece la validez de un instrumento de medición al comparar sus resultados con los de algún criterio externo que pretende medir lo mismo. Supongamos que Fernando trata de “medir” el grado en que es aceptado por Laura. Entonces decide que va a tomarla de la mano y observará su reacción. Supuestamente, si ella no retira la mano, esto indicaría cierta aceptación. Pero para asegurarse que su medición es válida, decide utilizar otra forma de medición adicional, por ejemplo, mirarla fijamente sin apartar la vista de sus ojos. En apariencia, si Laura le sostiene la mirada, esto sería otro indicador de aceptación. Así, su medición de aceptación se valida mediante dos métodos al comparar dos criterios. El ejemplo tal vez sea simple, pero describe la esencia de la validez relativa al criterio.

Este criterio es un estándar con el que se juzga la validez del instrumento (Wiersma y Jurs, 2008). Cuanto más se relacionen los resultados del instrumento de medición con el criterio, la validez de criterio será mayor. Por ejemplo, un investigador valida un examen sobre manejo de aviones al mostrar la exactitud con la que el examen predice qué tan bien un grupo de pilotos es capaz de operar un aeroplano.

Si el criterio se fija en el presente de manera paralela, se habla de **validez concurrente** (los resultados del instrumento se correlacionan con el criterio en el mismo momento o punto de tiempo). Por ejemplo, Núñez (2001) desarrolló una herramienta para medir el sentido de vida de acuerdo con la visión de Viktor Frankl, el test Celaya. Para aportar evidencia de validez en relación con su instrumento, lo aplicó y a su vez administró otros instrumentos que miden conceptos parecidos, tal como el PIL (Prueba de Propósito Vital) de Crumbaugh y Maholick (1969) y el Logo Test de Lukas (1984). Posteriormente comparó las puntuaciones de los participantes en las tres pruebas, demostró que las correlaciones entre las puntuaciones eran significativamente elevadas, de esta manera fue como aportó validez concurrente para su instrumento.

Si el criterio se fija en el futuro, se habla de **validez predictiva**. Por ejemplo, una prueba para determinar la capacidad gerencial de candidatos a ocupar altos puestos ejecutivos se validaría compa-

rando sus resultados con el desempeño posterior de los ejecutivos en su trabajo regular. Un cuestionario para detectar las preferencias del electorado por los distintos partidos contendientes y por sus candidatos en la época de las campañas, puede validarse comparando sus resultados con los resultados finales y definitivos de la elección.

El principio de la validez de criterio es sencillo: si diferentes instrumentos o criterios miden el mismo concepto o variable, deben arrojar resultados similares. Bostwick y Kyte (2005) lo expresan de la siguiente forma:

Si hay validez de criterio, las puntuaciones obtenidas por ciertos individuos en un instrumento deben estar correlacionadas y predecir las puntuaciones de estas mismas personas logradas en otro criterio.

La pregunta que se responde con la validez de criterio es: *¿en qué grado el instrumento comparado con otros criterios externos mide lo mismo?, o ¿qué tan cercanamente las puntuaciones del instrumento se relacionan con otro(s) resultado(s) sobre el mismo concepto?*

3. Evidencia relacionada con el constructo

La **validez de constructo** es probablemente la más importante, sobre todo desde una perspectiva científica, y se refiere a qué tan exitosamente un instrumento representa y mide un concepto teórico (Grinnell, Williams y Unrau, 2009). A esta validez le concierne en particular el significado del instrumento, esto es, qué está midiendo y cómo opera para medirlo. Integra la evidencia que soporta la interpretación del sentido que poseen las puntuaciones del instrumento (Messick, 1995).

Parte del grado en el que las mediciones del concepto proporcionadas por el instrumento se relacionan de manera consistente con otras mediciones de otros conceptos, de acuerdo con modelos e hipótesis derivadas teóricamente (que conciernen a los conceptos que se están midiendo) (Carmines y Zeller, 1991). A tales conceptos se les denomina constructos. Un **constructo** es una variable medida y que tiene lugar dentro de una hipótesis, teoría o un esquema teórico. Es un atributo que no existe aislado sino en relación con otros. No se puede ver, sentir, tocar o escuchar; pero debe ser inferido de la evidencia que tenemos en nuestras manos y que proviene de las puntuaciones del instrumento que se utiliza.

La validez de constructo incluye tres etapas (Carmines y Zeller, 1991):

1. Se establece y especifica la relación teórica entre los conceptos (sobre la base de la revisión de la literatura).
2. Se correlacionan los conceptos y se analiza cuidadosamente la correlación.
3. Se interpreta la evidencia empírica de acuerdo con el nivel en el 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 conveniente llevar a cabo tal validación, a menos que exista un marco teórico que soporte la variable en relación con otras variables. Desde luego, no es necesaria una teoría muy desarrollada, pero sí investigaciones que hayan demostrado que los conceptos se relacionan. Cuanto más elaborada y comprobada se encuentre la teoría que apoya la hipótesis, la validación del constructo arrojará mayor luz sobre la validez general de un instrumento de medición. Tenemos mayor confianza en la validez de constructo de una medición cuando sus resultados se correlacionan significativamente con un mayor número de mediciones de variables que, en teoría y de acuerdo con estudios antecedentes, están relacionadas. Veamos la validez de constructo con el ejemplo ya comentado sobre el clima organizacional.

Evidencia sobre la validez de constructo Debe explicar el modelo teórico empírico que subyace a la variable de interés.

Constructo Variable medida que tiene lugar dentro de una hipótesis, teoría o esquema teórico.

EJEMPLO

Hernández Sampieri (2005) aplicó un instrumento para evaluar al clima organizacional, el cual recordemos que midió 12 variables: (moral, apoyo de la dirección, innovación, etc.). La pregunta obvia es: ¿tal

instrumento realmente mide el clima organizacional?, ¿verdaderamente lo representa? En cuanto a contenido se demostró que sí reflejaba las principales dimensiones del clima organizacional. Pero esto no es suficiente, necesita demostrar que su instrumento es consistente con la teoría. Ésta, basada en diversos estudios, indica que tales dimensiones se encuentran fuertemente vinculadas y que se unen o funden entre sí para formar un constructo multidimensional denominado clima organizacional, y que además se asocian con el involucramiento en el trabajo y el compromiso organizacional. Entonces, para aportar validez de constructo, se correlacionaron todas las dimensiones entre sí y luego, la escala de clima con dicho involucramiento y compromiso. Tales vínculos se encontraron mediante análisis estadístico y los resultados coincidieron con la teoría y se obtuvo evidencia sobre la validez de constructo del instrumento.

Las preguntas que se responden con la validez de constructo son: *¿el concepto teórico está realmente reflejado en el instrumento?, ¿qué significan las puntuaciones del instrumento?, ¿el instrumento mide el constructo y sus dimensiones?, ¿por qué sí o por qué no?, ¿cómo opera el instrumento?*

Validez de expertos Se refiere al grado en que aparentemente un instrumento de medición mide la variable en cuestión, de acuerdo con expertos en el tema.

Otro tipo de validez que algunos autores consideran es la **validez de expertos** o *face validity*, la cual se refiere al grado en que aparentemente un instrumento de medición mide la variable en cuestión, de acuerdo con “voces calificadas”. Se encuentra vinculada a la validez de contenido y, de hecho, se consideró por muchos años como parte de ésta. Hoy se concibe como un tipo de evidencia distinta (Streiner y Norman, 2008). Regularmente se establece mediante la evaluación del instrumento ante expertos. Por ejemplo, Hernández Sampieri (2005) sometió el instrumento a revisión por parte de asesores en desarrollo organizacional, académicos y gerentes de recursos humanos. Asimismo, más recientemente se ha comentado en torno a la *validez consecuyente*, que se refiere a las secuelas sociales del uso e interpretación de una prueba (Mertens, 2005).

La validez total

La validez de un instrumento de medición se evalúa sobre la base de todos los tipos de evidencia. Cuanta mayor evidencia de validez de contenido, de validez de criterio y de validez de constructo tenga un instrumento de medición, éste se acercará más a representar la(s) variable(s) que pretende medir.

$$\text{Validez total} = \text{validez de contenido} + \text{validez de criterio} + \text{validez de constructo}$$

La relación entre la confiabilidad y la validez

Un instrumento de medición puede ser confiable, pero no necesariamente válido (un aparato, por ejemplo, quizá sea consistente en los resultados que produce, pero puede no medir lo que pretende).

La validez y la confiabilidad No se asumen, se prueban.

Por ello es requisito que el instrumento de medición demuestre ser *confiable y válido*. De no ser así, los resultados de la investigación no deben tomarse en serio.

Para ampliar este comentario, recurriremos a una analogía de Bostwick y Kyte (2005, pp. 108-109). Supongamos que vamos a probar un arma con tres tiradores. Cada uno debe realizar cinco disparos, entonces:

- Tirador 1 Sus disparos no impactan en el centro del blanco y se encuentran diseminados por todo el blanco.
- Tirador 2 Tampoco impacta en el centro del blanco, aunque sus disparos se encuentran cercanos entre sí, fue consistente, mantuvo un patrón.
- Tirador 3 Los disparos se encuentran cercanos entre sí e impactaron en el centro del blanco.

Sus resultados podrían visualizarse como en la figura 9.3, en la cual se vinculan la confiabilidad y la validez.



Figura 9.3 Representación de la confiabilidad y la validez.

Factores que pueden afectar la confiabilidad y la validez

Hay diversos factores que llegan a afectar la confiabilidad y la validez de los instrumentos de medición e introducen errores en la medición,¹ a continuación mencionaremos los más comunes.

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, que no debieran existir en la investigación.

También a las y los investigadores experimentados les toma cierto tiempo desarrollar un instrumento de medición. Además, para construir un instrumento de medición se requiere conocer muy bien la variable que se pretende medir, así como la teoría que la sustenta.

El segundo factor es que a veces se utilizan instrumentos desarrollados en el extranjero que no han sido validados en nuestro contexto: cultura y tiempo. Traducir un instrumento, aun cuando adaptemos los términos a nuestro lenguaje y los contextualicemos, no es ni remotamente una validación. Es un primer y necesario paso, aunque sólo es el principio. En el caso de traducciones, es importante verificar que los términos centrales tengan referentes con el mismo significado —o alguno muy parecido— en la cultura en la que se va a utilizar dicho instrumento (vincular términos entre la cultura de origen y la cultura destinataria). A veces se traduce, se obtiene una versión y ésta, a su vez, se vuelve a traducir de nuevo al idioma original.

Por otra parte, existen instrumentos que fueron validados en nuestro contexto, pero hace mucho tiempo. Hay instrumentos en los que hasta el lenguaje nos suena “anticuado”. Las culturas, los grupos y las personas cambian; y esto debemos tomarlo en cuenta al elegir o desarrollar un instrumento de medición.

Un tercer factor es que en ocasiones el instrumento resulta inadecuado para las personas a quienes se les aplica: no es empático. Utilizar un lenguaje muy elevado para el sujeto participante, no tomar en cuenta diferencias en cuanto a género, edad, conocimientos, memoria, nivel ocupacional y educativo, motivación para contestar, capacidades de conceptualización y otras diferencias en los participantes, son errores que llegan a afectar la validez y confiabilidad del instrumento de medición. Este error ocurre a menudo cuando los instrumentos deben aplicarse a niños. Asimismo, hay grupos de la pobla-

¹ Se ha omitido intencionalmente la exposición de los errores sistemáticos y no sistemáticos que afectan a la confiabilidad y la validez, con objeto de simplificar al lector las explicaciones. Un comentario se incluye en el CD anexo (Material complementario → Capítulos → Capítulo 7 “Recolección de los datos cuantitativos: segunda parte”).



ción que requieren instrumentos apropiados para ellos, tal es el caso de las personas con capacidades distintas. En la actualidad se han desarrollado diversas pruebas que las toman en cuenta (por ejemplo, pruebas en sistema Braille para personas con discapacidades visuales o pruebas orales para individuos que no pueden escribir). Otro ejemplo lo son los indígenas o inmigrantes de otras culturas, en ocasiones se les administran instrumentos que no toman en cuenta su lenguaje y contexto.

Quien realiza una investigación debe siempre adaptarse a los participantes y no éstos a él o ella, ya que es necesario brindarles todo tipo de facilidades. Si éste es el caso, sugerimos que se consulte a Mertens y McLaughlen (2004), en cuyo libro tienen un capítulo dedicado a la recolección de información de personas con capacidades diferentes o de culturas especiales, y a Eckhardt y Anastas (2007). Asimismo, es recomendable revisar la página web de alguna asociación internacional como la American Psychological Association.

El cuarto factor agrupa diversas cuestiones vinculadas con los estilos personales de los participantes (Bostwick y Kyte, 2005) tales como: discapacidad social (tratar de dar una impresión muy favorable a través de las respuestas), tendencia a asentir con respecto a todo lo que se pregunta, dar respuestas inusuales o contestar siempre negativamente.

Un quinto factor que puede influir *está constituido por las condiciones en las que se aplica el instrumento de medición*. El ruido, la iluminación, el frío (por ejemplo, en una encuesta de casa en casa), un instrumento demasiado largo o tedioso, una encuesta telefónica después de que algunas compañías han utilizado el mercadeo telefónico en exceso y a destiempo (promocionar servicios a las 7 a.m. de un domingo o después de las 11 p.m. entre semana) son cuestiones que llegan a afectar negativamente la validez y la confiabilidad, al igual que si el tiempo que se brinda para responder al instrumento es inapropiado. Por lo común en los experimentos se cuenta con instrumentos de medición más largos y complejos que en los diseños no experimentales. Por ejemplo, en una encuesta pública sería muy difícil aplicar una prueba larga o compleja.

El sexto elemento es la falta de estandarización. Por ejemplo, que las instrucciones no sean las mismas para todos los participantes, que el orden de las preguntas sea distinto para algunos individuos, que los instrumentos de observación no sean equivalentes, etc. Este elemento también se vincula con la objetividad.

Aspectos mecánicos tales como que si el instrumento es escrito, que no sean legibles las instrucciones, falten páginas, no haya espacio adecuado para contestar o no se comprendan las instrucciones, también influyen de manera desfavorable.

Con respecto a la validez de constructo dos factores pueden afectarla significativamente: *a)* la estrechez del contenido, es decir que se excluyan dimensiones importantes de la variable o las variables medidas y *b)* la amplitud exagerada, donde el riesgo es que el instrumento contenga excesiva intrusión de otros constructos.

Muchos de los errores se pueden evitar mediante una adecuada revisión de la literatura, que nos permite seleccionar las dimensiones apropiadas de las variables del estudio, criterios para comparar los resultados de nuestro instrumento, teorías de respaldo, instrumentos de dónde elegir, etcétera.

La objetividad

Se trata de un concepto difícil de lograr, particularmente en el caso de las ciencias sociales. En ciertas ocasiones se alcanza mediante el consenso (Grinnell, Williams y Unrau, 2009). Al tratarse de cuestiones físicas las percepciones suelen compartirse (por ejemplo, la mayoría de las personas estarían de acuerdo en que el agua de mar contiene sal o los rayos del Sol queman), pero en temas que tienen que ver con la conducta humana como los valores, las atribuciones y las emociones, el consenso es más complejo. Imaginemos que 10 observadores deben ver una película y calificarla como “muy violenta”, “violenta”, “neutral”, “poco violenta” y “nada violenta”. Tres personas indican que es muy violenta, tres que es violenta y cuatro la evalúan como neutral; qué tan violenta es la película resulta un cuestionamiento difícil. O bien, ¿quién fue mejor compositor: Mozart, Beethoven o Bach? Todo es relativo. Sin embargo, la objetividad aumenta al reducirse la incertidumbre (Unrau, Grinnell y Williams, 2005).

Desde luego, la certidumbre total no existe ni en las ciencias físicas; el conocimiento es aceptado como verdadero, hasta que nueva evidencia demuestra lo contrario.

En un instrumento de medición, la **objetividad** se refiere al grado en que éste es permeable a la influencia de los sesgos y tendencias del investigador o investigadores que lo administran, califican e interpretan (Mertens, 2005). Investigadores racistas o “machistas” quizás influyan negativamente por su sesgo contra un grupo étnico o el género femenino. Lo mismo podría suceder con las tendencias ideológicas, políticas, religiosas o la orientación sexual. En este sentido, los aparatos y sistemas calibrados (por ejemplo, una pistola láser para medir la velocidad de un automóvil) son más objetivos que otros sistemas que requieren cierta interpretación (como un detector de mentiras) y éstos, a su vez, más objetivos que las pruebas estandarizadas, las cuales son menos subjetivas que las pruebas proyectivas.

La objetividad se refuerza mediante la estandarización en la aplicación del instrumento (mismas instrucciones y condiciones para todos los participantes) y en la evaluación de los resultados; así como al emplear personal capacitado y experimentado en el instrumento. Por ejemplo, si se utilizan observadores, su proceder en todos los casos debe ser lo más similar que sea posible y su entrenamiento tendrá que ser profundo y adecuado.

Los estudios cuantitativos buscan que la influencia de las características y las tendencias del investigador se reduzca al mínimo posible, lo que insistimos es un ideal, pues la investigación siempre es realizada por seres humanos.

La validez, la confiabilidad y la objetividad no deben tratarse de forma separada. Sin alguna de las tres, el instrumento no es útil para llevar a cabo un estudio.

Objetividad del instrumento Se refiere al grado en que el instrumento es permeable a la influencia de los sesgos y tendencias de los investigadores que lo administran, califican e interpretan.

¿Cómo se sabe si un instrumento de medición es confiable y válido?

En la práctica es casi imposible que una medición sea perfecta. Generalmente se tiene un grado de error. Desde luego, se trata de que este error sea el mínimo posible, por lo cual la medición de cualquier fenómeno se conceptualiza con la siguiente fórmula básica:

$$X = t + e$$

Donde X representa los valores observados (resultados disponibles); t , los valores verdaderos; y e , el grado de error en la medición. Si no hay un error de medición (e es igual a cero), el valor observado y el verdadero son equivalentes. Esto puede verse claramente así:

$$\begin{aligned} X &= t + 0 \\ X &= t \end{aligned}$$

Esta situación representa el ideal de la medición. Cuanto 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 la medición está contaminada por un grado de error considerable, la motivación registrada por el instrumento será bastante diferente de la motivación real de ese individuo. Por ello, es importante que el error se reduzca lo más posible. Pero, ¿cómo sabemos el grado de error que tenemos en una medición? Al calcular la confiabilidad y la validez.

Cálculo de la confiabilidad o fiabilidad

Existen diversos procedimientos para calcular la confiabilidad de un instrumento de medición. Todos utilizan procedimientos y fórmulas que producen coeficientes de fiabilidad. La mayoría de éstos pueden oscilar entre cero y uno, donde un coeficiente de cero significa nula confiabilidad y uno representa un máximo de confiabilidad (fiabilidad total, perfecta). Cuanto más se acerque el coeficiente a cero (0), mayor error habrá en la medición. Esto se ilustra en la figura 9.4.



Figura 9.4 Interpretación de un coeficiente de confiabilidad.

Los procedimientos más utilizados para determinar la confiabilidad mediante un coeficiente son: 1) *medida de estabilidad* (confiabilidad por *test-retest*), 2) *método de formas alternativas o paralelas*, 3) *método de mitades partidas* (*split-halves*) y 4) *medidas de consistencia interna*. Estos procedimientos no se detallan en esta sección, sino que se explican en el capítulo 10, “Análisis de los datos cuantitativos”, debido a que requieren del entendimiento de ciertos conceptos estadísticos. Simplemente comentaremos su interpretación con la medida de consistencia interna denominada “coeficiente alfa Cronbach”, que tal vez es la más utilizada.

Supongamos que una investigadora desarrolló un instrumento para medir el grado de “amor romántico” entre parejas de jóvenes universitarios, el cual se fundamentó en cuatro de las herramientas más conocidas para ello: la medida de Rubin sobre el amar y el vincularse con los demás, la escala sobre actitudes hacia el amor, la medida sobre el amor apasionado y la escala del amor triangular (Graham y Christiansen, 2009). Para estimar la confiabilidad de su instrumento lo debe aplicar a su muestra y sobre la base de los resultados calcular tal coeficiente. Imaginemos que obtiene un valor alfa Cronbach de 0.96, que es muy elevado, lo que significa que su medida del “amor romántico” es sumamente confiable, esto se representa en la figura 9.5.

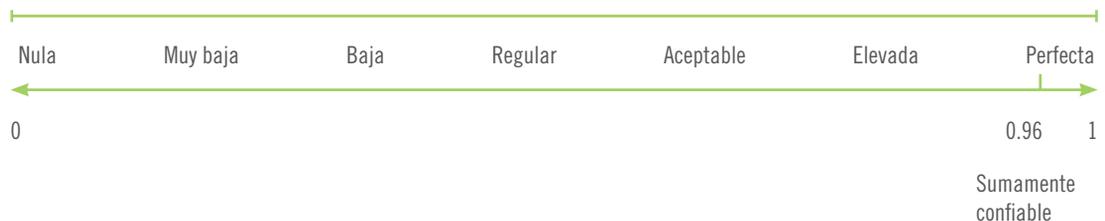


Figura 9.5 Interpretación de un coeficiente de confiabilidad sobre un instrumento que mide el “amor romántico”.

La confiabilidad varía de acuerdo con el número de ítems² que incluya el instrumento de medición. Cuantos más ítems haya, mayor será ésta. Lo cual resulta lógico; veámoslo con un ejemplo cotidiano: si se desea probar qué tan confiable o consistente es la lealtad de un amigo hacia nuestra persona, cuantas más pruebas le pongamos, su fiabilidad será mayor. Claro está que demasiados ítems provocarán cansancio en los participantes.

Cada vez que se administra un instrumento de medición debe calcularse la confiabilidad, al igual que evaluarse la evidencia sobre la validez.

² Un ítem es la unidad mínima que compone una medición; es un reactivo que estimula una respuesta en un sujeto (por ejemplo, una pregunta, una frase, una lámina, una fotografía o un objeto de descripción).

Cálculo de la validez

Con respecto a la validez de contenido, primero es necesario revisar cómo ha sido medida la variable por otros investigadores. Y, con base en dicha revisión, elaborar un universo de ítems o reactivos posibles para medir la variable y sus dimensiones (el universo debe ser lo más exhaustivo que sea posible). Después, se consulta a investigadores familiarizados con la variable para ver si el universo es verdaderamente exhaustivo. Se seleccionan los ítems bajo una cuidadosa evaluación, uno por uno. Y si la variable está compuesta por diversas dimensiones o facetas, se extrae una muestra probabilística de reactivos, ya sea al azar o estratificada (cada dimensión constituiría un estrato). Se administran los ítems, se correlacionan las puntuaciones de éstos entre sí (tiene que haber correlaciones altas, en especial entre ítems que miden una misma dimensión, pero teniendo cuidado que sean capaces de discriminar entre participantes) (Bohrnstedt, 1976; Punch, 2009); y se hacen estimaciones estadísticas para ver si la muestra es representativa. Para calcular la validez de contenido son necesarios varios coeficientes. Éste sería un procedimiento ideal. Pero, como veremos más adelante, a veces no se calculan estos coeficientes, sino que se seleccionan los ítems mediante un proceso que asegura la representatividad (no de manera estadística sino conceptual).

La *validez de criterio* se estima al correlacionar la medición con el criterio externo (puntuaciones del instrumento frente a las puntuaciones en el criterio), y este coeficiente se toma como coeficiente de validez (Bohrnstedt, 1976). Que podría representarse con el ejemplo de la figura 9.6.³



Figura 9.6 Ejemplo para la estimación de la validez de criterio.

La *validez de constructo* suele determinarse mediante procedimientos de análisis estadístico multivariado (“análisis de factores”, “análisis discriminante”, “regresiones múltiples”, etc.), los cuales se revisan en el CD anexo: Material complementario → Capítulos → Capítulo 8: “Análisis estadístico: segunda parte”.



¿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 y aplicarlos es semejante. Éste se resume mediante etapas en el diagrama de la figura 9.7 y corresponde a la parte del plan de recolección que contesta a la pregunta: ¿a través de qué método vamos a recolectar los datos? Y cabe señalar que cada etapa o fase no se detalla en este capítulo, sino en el CD anexo: Material complementario → Capítulos → Capítulo 7: “Recolección de los datos cuantitativos: segunda parte”.

OQ5



³ Las pruebas de correlación se presentan en el siguiente capítulo: “Análisis de los datos cuantitativos”.



Figura 9.7 Proceso para construir un instrumento de medición.

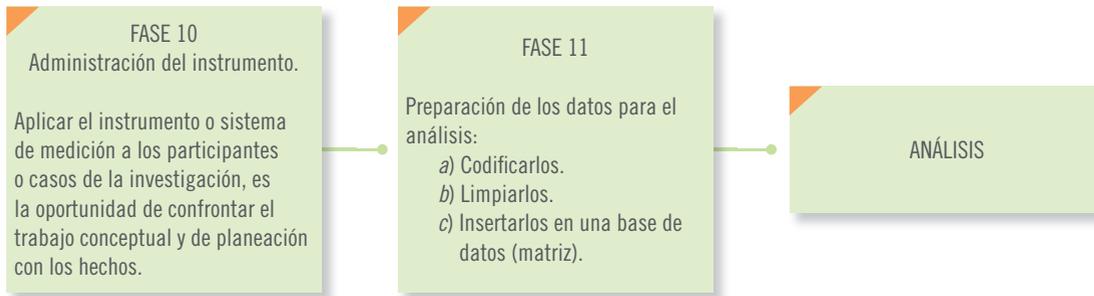


Figura 9.7 Proceso para construir un instrumento de medición (continuación).

Las fases 1 a 7 del diagrama se refieren propiamente al desarrollo del instrumento o sistema de medición, mientras que las etapas 8 a 11 representan la administración del mismo y la preparación de los datos para su análisis.

Tres cuestiones fundamentales para un instrumento o sistema de medición

Existen tres cuestiones básicas a considerar al momento de construir un instrumento.

El tránsito de la variable al ítem

Cuando se construye un instrumento, el proceso más lógico para hacerlo es transitar de la variable a sus dimensiones o componentes, luego a los indicadores y finalmente a los ítems o reactivos. En la tabla 9.2 podemos ver ejemplos de este tránsito.

Tabla 9.2 Ejemplo de desarrollo de ítems

Estudio de las preferencias de los jóvenes para divertirse (ejemplo sencillo) ⁴			
Variable	Dimensión	Indicadores	Ítems
Preferencia de actividad para salir con alguien del género opuesto.	Actividad nocturna entre semana.	Jerarquía de preferencias de actividades de lunes a jueves (aunque algunos comienzan el fin de semana desde el jueves).	De lunes a jueves, ¿cuál sería tu actividad preferida nocturna para salir con el chico o chica que más te gusta? (marcar la que más te agrada). 1. Salir a cenar a un restaurante. 2. Ir al cine. 3. Ir a un bar, “antro”, <i>grill</i> , etcétera. 4. Acudir a una taberna o cervecería. 5. Ir a bailar a una discoteca, disco o “antro”. 6. Ir a una fiesta privada. 7. Acudir al teatro. 8. Acudir a un concierto. 9. Pasear por un parque, jardín o avenida. 10. Otra (especificar).
	Actividad nocturna en fin de semana.	Jerarquía de preferencias de actividades en viernes y sábado.	Mismas categorías u opciones de respuesta.
	Actividad nocturna en domingo.	Jerarquía de preferencias de actividades en domingo.	Mismas categorías y opciones de respuesta.

(continúa)

⁴ En los ejemplos de la tabla sólo se incluyen unos cuantos ítems o reactivos por cuestiones de espacio, son ejemplos muy resumidos.

▲ **Tabla 9.2** Ejemplo de desarrollo de ítems. (continuación)

Investigación sobre el clima organizacional			
Variable	Dimensión	Indicadores	Ítems
Clima organizacional.	Moral.	Grado en que los miembros de una organización o departamento perciben que colaboran y cooperan entre sí, se apoyan mutuamente y mantienen relaciones de amistad y compañerismo.	<ul style="list-style-type: none"> • Mis compañeros de trabajo son mis amigos 5. Totalmente de acuerdo. 4. De acuerdo. 3. Ni de acuerdo ni en desacuerdo. 2. En desacuerdo. 1. Totalmente en desacuerdo. <ul style="list-style-type: none"> • En mi trabajo hay mucho compañerismo. (Mismas opciones de respuesta que el ítem anterior.) • Siempre que lo necesito mis compañeros de trabajo me brindan apoyo. (Mismas opciones de respuesta.) • En el departamento donde trabajo nos mantenemos unidos. (Mismas opciones de respuesta.) • La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros. (Mismas opciones de respuesta.) • ¿Qué tanto apoyo le brindan sus compañeros cuando usted lo necesita? 5. Total. 4. Bastante. 3. Aceptablemente. 2. Poco. 1. Ninguno.
	Autonomía.	Grado de libertad percibida para tomar decisiones y realizar el trabajo.	<ul style="list-style-type: none"> • En esta empresa tengo libertad para tomar decisiones que tienen que ver con mi trabajo. 5. Totalmente de acuerdo. 4. De acuerdo. 3. Ni de acuerdo ni en desacuerdo. 2. En desacuerdo. 1. Totalmente en desacuerdo. <ul style="list-style-type: none"> • Mi jefe me da libertad para tomar decisiones que tienen que ver con mi trabajo. (Mismas opciones de respuesta que el ítem anterior.)
	Atribución del desempeño.	Grado de conciencia compartida por desempeñarse con calidad en las tareas laborales, sobre la base de la cooperación.	<ul style="list-style-type: none"> • En esta empresa todos tratamos de hacer bien nuestro trabajo. (Mismas opciones de respuesta que el ítem anterior.) • En esta empresa todos queremos dar lo mejor de nosotros en el trabajo. (Mismas opciones.)

Codificación

Codificar los datos significa asignarles un valor numérico o símbolo que los represente. Es decir, a las categorías (opciones de respuesta o valores) de cada ítem y variable se les asignan valores numéricos o signos que tienen un significado. Por ejemplo, si tuviéramos la variable “género” con sus respectivas categorías, masculino y femenino, a cada categoría le asignaríamos un valor. Esto podría ser:

Codificación Significa asignar a los datos un valor numérico o símbolo que los represente, ya que es necesario para analizarlos cuantitativamente.

Categoría	Codificación (valor asignado)
Masculino	1
Femenino	2

Así, Paola Yáñez en la variable género sería 2, Luis Gerardo Vera y José Ramón Calderón serían 1, Liz Rangel 2 y así sucesivamente.

Otro ejemplo sería la variable “horas de exposición diaria a la televisión”, que podría codificarse como se muestra en la tabla 9.3.

▲ **Tabla 9.3** Ejemplo de codificación

Categoría	Codificación (valor asignado)
— No ve televisión	0
— Menos de una hora	1
— Una hora	2
— Más de una hora, pero menos de dos	3
— Dos horas	4
— Más de dos horas, pero menos de tres	5
— Tres horas	6
— Más de tres horas, pero menos de cuatro	7
— Cuatro horas	8
— Más de cuatro horas	9

En el primer ejemplo de la tabla 9.2, la respuesta a la pregunta: *de lunes a jueves, ¿cuál sería tu actividad nocturna preferida para salir con el chico o chica que más te gusta?*, la codificación era con números (1 = salir a cenar a un restaurante; 2 = ir al cine; 3 = ir a un bar, “antro”, grill, etc.; 4 = acudir a una taberna o cervecería, 5 = ir a bailar a una discoteca, disco o “antro”; 6 = ir a una fiesta privada; 7 = acudir al teatro; 8 = acudir a un concierto; 9 = pasear por un parque, jardín, avenida, y 10 = otra).

Mientras que en el ítem: “En esta empresa tengo libertad para tomar decisiones que tienen que ver con mi trabajo”, la codificación era:

5. Totalmente de acuerdo.
4. De acuerdo.
3. Ni de acuerdo ni en desacuerdo.
2. En desacuerdo.
1. Totalmente en desacuerdo.

Es necesario insistir que cada ítem y variable deberán tener una codificación (códigos numéricos o simbólicos) para sus categorías, a esto se le conoce como “precodificación”. Desde luego, hay veces que un ítem no puede ser codificado *a priori* (precodificado), porque es muy 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 cuántas y cuáles serían. En tales situaciones, la codificación se lleva a cabo una vez que se aplica el ítem (*a posteriori*). Éste es el caso de algunos ítems que por ahora denominaremos “abiertos”.

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). La codificación puede o no incluirse en el instrumento de medición, veámoslo con un ejemplo de pregunta:

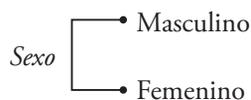
Pregunta precodificada	Pregunta no precodificada
¿Tiene usted novia?	¿Tiene usted novia?
<input type="checkbox"/> 1 Sí	<input type="checkbox"/> Sí
<input type="checkbox"/> 0 No	<input type="checkbox"/> No

Asimismo, es muy importante indicar el nivel de medición de cada ítem y, por ende, el de las variables, porque es parte de la codificación y 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 de la prueba para correlacionar dos variables ordinales). Así, es necesario hacer una relación de variables, ítems y niveles de medición.

Niveles de medición

Existen cuatro niveles de medición ampliamente conocidos.

1. *Nivel de medición nominal.* En este nivel hay dos o más categorías del ítem o la variable. Las categorías no tienen orden ni jerarquía. Lo que se mide (objeto, persona, etc.) se coloca en una u otra categorías, lo cual indica tan sólo diferencias respecto de una o más características. Por ejemplo, la variable “género” de la persona posee sólo dos categorías: masculino y femenino. Ninguna de las categorías implica 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 un símbolo a cada categoría, esto identificará exclusivamente a la categoría. Por ejemplo:

* = Masculino
z = Femenino

Si usamos numerales, es lo mismo:

1 = Masculino 2 = Masculino

es igual a

2 = Femenino 1 = Femenino

Los números utilizados en este nivel de medición tienen una función puramente de clasificación y *no* se pueden manipular de manera aritmética. Por ejemplo, la afiliación religiosa es una variable nominal; si pretendiéramos operarla de forma aritmética se presentarían situaciones tan ridículas como ésta:

- 1** = Católico
2 = Judío **1 + 2 = 3**
3 = Protestante
4 = Musulmán ¿Un católico + un judío = un protestante?
5 = Otros (No tiene sentido)

Las variables nominales pueden incluir dos categorías (dicotómicas), o bien, tres o más categorías (categóricas). Ejemplos de variables nominales dicotómicas serían el género, el veredicto de un jurado (culpable-no culpable) y el tipo de escuela a la que se asiste (privada-pública); y como ejemplos de variables nominales categóricas tendríamos la afiliación política (partido A, partido B, etc.), la licenciatura elegida, el grupo étnico, el departamento, la provincia o el estado de nacimiento, la clase de material de construcción (“no” su resistencia, ésta sería otra variable), tipo de medicamento suministrado (“no” la dosis, que sería una variable distinta), bloques de mercado (asiático, latinoamericano, comunidad europea, etc.) y el canal de televisión preferido.

2. *Nivel de medición ordinal.* En este nivel hay varias categorías, pero además mantienen un orden de mayor a menor. Las etiquetas o los símbolos de las categorías sí indican jerarquía. Por ejemplo, el prestigio ocupacional en Estados Unidos se ha medido por diversas escalas que reordenan las profesiones de acuerdo con su prestigio, por ejemplo:⁵

Valor en escala	Profesión
90	Ingeniero químico
80	Científico de ciencias naturales (excluyendo la química)
60	Actor común
50	Operador de estaciones eléctricas de potencia
02	Manufacturero de tabaco

Los números (símbolos de categorías) definen posiciones, en el ejemplo: 90 es más que 80, 80 más que 60, 60 más que 50 y así sucesivamente. 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 eléctricas (50) existe la misma distancia en prestigio que entre un científico de ciencias naturales (80) y un ingeniero químico (90). Al parecer, en ambos casos la distancia es 10, pero no es una distancia real. Otra escala⁶ clasificó el prestigio de dichas profesiones de la siguiente manera:

Valor en escala	Profesión
98	Ingeniero químico
95	Científico de ciencias naturales (excluyendo la química)
84	Actor común
78	Operador de estaciones eléctricas de potencia
13	Manufacturero de tabaco

Aquí la distancia entre un actor (84) y un operador de estaciones (78) es de seis, y la distancia entre un ingeniero químico (98) y un científico de ciencias naturales (95) es de tres. Otro ejemplo sería la posición jerárquica en la empresa:

⁵ Duncan (1977).

⁶ Nam *et al.* (1965) y Nam (1983).

Presidente	10
Vicepresidente	9
Director general	8
Gerente de área	7
Subgerente o superintendente	6
Jefe	5
Empleado <i>A</i>	4
Empleado <i>B</i>	3
Empleado <i>C</i>	2
Intendencia	1

Sabemos que el presidente (10) es más que el vicepresidente (9), éste más que el director general (8), a su vez este último más que el gerente (7) y así sucesivamente; pero no se precisa en cada caso cuánto más. Tampoco se utilizan 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. Otros ejemplos de este nivel serían: la medición por rangos de las preferencias de marcas de bebidas refrescantes con gas (refrescos o sodas), autopercepción del grado de dolor de cabeza y jerarquización de valores (en primer lugar, en segundo lugar, en tercero).

3. *Nivel de medición por intervalos.* Además del orden o la jerarquía entre categorías, se establecen intervalos iguales en la medición. Las distancias entre categorías son las mismas a lo largo de toda la escala, por lo que hay un intervalo constante, una unidad de medida.



Figura 9.8 Escala con intervalos iguales entre categorías.

Por ejemplo, en una prueba de resolución de problemas matemáticos (30 problemas de igual dificultad). Si Ana Cecilia resolvió 10, Laura resolvió 20 y Abigail, 30. La distancia entre Ana Cecilia y Laura es igual a la distancia entre Laura y Abigail.

Sin embargo, el cero (0) en la medición es un cero arbitrario, no es real, ya que 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, que puede medirse en grados centígrados y Fahrenheit: el cero es arbitrario, pues no implica que en realidad 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 (por ejemplo, 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 utilizarían. Aunque algunos investigadores no están de acuerdo con suponer tales mediciones como si fueran de intervalo. El producto interno bruto o producto nacional bruto estaría en este estadio.

4. *Nivel de medición de razón.* En este nivel, además de tenerse todas las características del nivel de intervalos (periodos iguales entre las categorías, y aplicación de operaciones aritméticas básicas y sus derivaciones), el cero es real y es absoluto (no es arbitrario). Cero absoluto implica que hay un punto en la escala donde está ausente o no existe la propiedad medida (vea la figura 9.9).



El cero es real

Figura 9.9 Ejemplo de escala para el nivel de medición de razón.

Ejemplos de estas mediciones serían la exposición a la televisión (en minutos), el número de hijos, las ventas de un producto, los metros cuadrados de construcción, ingresos (en moneda), presión arterial, etcétera.

¿De qué tipos de instrumentos de medición o recolección de datos cuantitativos disponemos en la investigación?

En la investigación disponemos de diversos tipos de instrumentos para medir las variables de interés y en algunos casos llegan a combinarse varias técnicas de recolección de los datos. A continuación las describimos brevemente.

Los instrumentos que serán revisados en este capítulo son: cuestionarios⁷ y escalas de actitudes. En el capítulo “Recolección de los datos cuantitativos, segunda parte”, que encontrará en el CD anexo, se comentarán los siguientes: registros del contenido (análisis de contenido) y observación, pruebas estandarizadas (medidas del desempeño individual), recolección de información factual e indicadores (análisis de datos secundarios de registros públicos y documentación) y metaanálisis, así como otras mediciones.

La codificación y la preparación de los datos obtenidos se discutirán después de presentar los principales instrumentos de medición.

OQ4



Cuestionarios

Tal vez el instrumento más utilizado para recolectar los datos es el cuestionario. Un **cuestionario** consiste en un conjunto de preguntas respecto de una o más variables a medir. Debe ser congruente con el planteamiento del problema e hipótesis (Brace, 2008). Comentaremos primero sobre las preguntas y luego sobre las características deseables de este tipo de instrumento, así como los contextos en los cuales se pueden administrar los cuestionarios.

OQ4

Cuestionario Tal vez sea el instrumento más utilizado para recolectar los datos, consiste en un conjunto de preguntas respecto de una o más variables a medir.

¿Qué tipos de preguntas se pueden hacer?

El contenido de las preguntas de un cuestionario es tan variado como los aspectos que mide. Básicamente se consideran *dos tipos de preguntas: cerradas y abiertas*.

OQ5

Preguntas cerradas

Las **preguntas cerradas** contienen categorías u opciones de respuesta que han sido previamente delimitadas. Es decir, se presentan las posibilidades de respuesta a los participantes, quienes deben acotarse a éstas. Pueden ser dicotómicas (dos posibilidades de respuesta) o incluir varias opciones de respuesta. Ejemplos de preguntas cerradas dicotómicas serían:

Preguntas cerradas Son aquellas que contienen opciones de respuesta previamente delimitadas. Son más fáciles de codificar y analizar.

- | | |
|-----------------------------|---|
| ¿Estudia usted actualmente? | ¿Durante la semana pasada vio la final de la Liga de Campeones de Europa? |
| () Sí | () Sí |
| () No | () No |

⁷ Las entrevistas se plantean como un contexto en el cual pueden ser administrados los cuestionarios.

Ejemplos de preguntas cerradas con varias opciones de respuesta serían:

Como usted sabe todos los países desarrollados reciben inmigrantes. ¿Cree que, en términos generales, la inmigración es positiva o más bien negativa para estos países?

- Positiva
- Ni positiva ni negativa
- Negativa
- No sabría decir

¿Cuál es el puesto que ocupa usted en su empresa?

- | | |
|---|--|
| <input type="checkbox"/> Presidente/Director general | <input type="checkbox"/> Vicepresidente/Director corporativo |
| <input type="checkbox"/> Subdirector/Director/Gerente | <input type="checkbox"/> Subgerente/Superintendente |
| <input type="checkbox"/> Coordinador | <input type="checkbox"/> Jefe de área |
| <input type="checkbox"/> Supervisor | <input type="checkbox"/> Empleado |
| <input type="checkbox"/> Obrero | <input type="checkbox"/> Otro (especificar) _____ |

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 muestran al encuestado, quien debe elegir la opción que describa más adecuadamente su respuesta. Gambaro (2002) hace notar algo muy lógico pero que en ocasiones se descuida y resulta fundamental: cuando las preguntas presentan varias opciones, éstas deben recoger todas las posibles respuestas.

Ahora bien, hay preguntas cerradas donde el participante puede seleccionar más de una opción o categoría de respuesta (*posible multirrespuesta*).

EJEMPLO

Supongamos que un entrevistador pregunta:

¿Esta familia tiene en el hogar...? (Marque con una cruz o tache todas las opciones que el entrevistado o entrevistada señale que tiene en su hogar):

- | | |
|--|--|
| <input type="checkbox"/> Radio | <input type="checkbox"/> Televisión |
| <input type="checkbox"/> Reproductor de DVD | <input type="checkbox"/> TV de paga (SKY, Cablevisión, DirectTV, otros sistemas locales de cable o TV satelital) |
| <input type="checkbox"/> Computadora | <input type="checkbox"/> Internet |
| <input type="checkbox"/> Teléfono celular o móvil | <input type="checkbox"/> Equipo de sonido para CD |
| <input type="checkbox"/> iPod | |
| <input type="checkbox"/> Teléfono (línea telefónica en casa) | |

En preguntas como la del ejemplo anterior, los participantes pueden marcar una, dos, tres, cuatro o más opciones de respuesta. Las categorías no son mutuamente excluyentes.

EJEMPLO

Otro ejemplo de esta clase de preguntas sería el siguiente:

De los siguientes servicios que presta la sala de lectura de la biblioteca, ¿cuál o cuáles utilizó el semestre anterior? (Puede señalar más de una opción.)

- No entré
- Consultar algún libro
- Consultar algún periódico
- Consultar alguna revista
- Estudiar
- Utilizar la computadora para buscar referencias y documentos en internet
- Utilizar la computadora para revisar mi correo electrónico
- Utilizar la computadora para elaborar un trabajo
- Buscar a alguna persona
- Otros (especificar)

En ocasiones, el encuestado tiene que jerarquizar opciones.

EJEMPLO

De las siguientes compañeras de clase, ¿quién te atrae más?, ¿cuál en segundo lugar?, ¿cuál en tercer lugar?, ¿cuál en cuarto lugar? y ¿cuál en quinto y último lugar?

- Sandra
- Lucía
- Ana
- Mariana
- Paola

O bien, en otras preguntas se debe designar un puntaje a una o diversas cuestiones.

EJEMPLO

A continuación voy a mencionarle algunos de los problemas que suelen preocupar a los habitantes de este municipio y le pediría que en cada caso me dijera: ¿qué tanto le preocupa a usted cada uno de ellos?; donde 10 significa: “me preocupa muchísimo” y 0 quiere decir: “no me preocupa en absoluto”.⁸

- _____ Desempleo
- _____ Pobreza
- _____ Inseguridad al transitar por la calle o viajar en transporte público
- _____ Empleo mal remunerado/bajos salarios
- _____ Robos/asaltos en los hogares y viviendas
- _____ Robos de vehículos/autos, motocicletas, bicicletas
- _____ Pandillerismo
- _____ Venta de drogas-narcomenudeo
- _____ Secuestros
- _____ Recolección de la basura (no todos los días la recogen)
- _____ Horarios inadecuados para la recolección de la basura
- _____ Escasez de agua
- _____ Cortes en el suministro de agua
- _____ Falta de vivienda
- _____ Servicios de salud insuficientes
- _____ Carencia/deficiencia de servicios educativos
- _____ Drenaje inadecuado en las calles
- _____ Tránsito/tráfico/vialidad
- _____ Pavimentación y bacheo mal hechos

⁸ Esta pregunta fue administrada a personas que tuvieran estudios mínimos de bachillerato o preparatoria, aunque funcionó con personas cuyo nivel era de secundaria. No se incluyeron todos los problemas por cuestiones de espacio, solamente algunos para ilustrar el tipo de pregunta.

- _____ Falta de infraestructura (calles, puentes, etcétera)
- _____ Corrupción de funcionarios municipales, policías, agentes de tránsito y vialidad
- _____ Situación económica familiar

En otras preguntas, se anota una cifra dentro de un rango predeterminado:

EJEMPLO

Aproximadamente, ¿cuántos minutos dedica diariamente a hacer deporte entre semana, es decir, de lunes a viernes?

En algunas más, el encuestado se ubica en una escala. El concepto de escala (aplicado a la medición) puede definirse como: “sucesión ordenada de valores distintos de una misma cualidad” (Real Academia Española, 2001, p. 949). Es un patrón, conjunto, medida o estimación regular de acuerdo con algún estándar o tasa, respecto de una variable. Ejemplos: escala de temperatura en grados centígrados, escala de inteligencia, escala de distancia en kilómetros, metros y centímetros; escala de peso en kilogramos, escala musical con octavas, etcétera.

EJEMPLO

¿Qué tan enamorada está usted de su novio? (Del 0 al 100)

☺	100 – Completamente enamorada	.	
	99	.	
	98	.	
	.	20	
	.	10	
	.	.	
	80	.	
	70	2	
	60	1	
☹	50	⊕	0 – Nada enamorada

Finalmente, en ocasiones se encadenan varias preguntas en una, como en el siguiente ejemplo (los candidatos y candidatas son ficticios(as)). Cualquier parecido con algún nombre es mera coincidencia.

EJEMPLO

Le voy a mencionar algunos nombres de políticos de nuestro municipio y le pediría que en cada caso me dijera si sabe usted quién es y a qué partido pertenece, así como su opinión de esta persona:

Político(a) (Rotar opciones)	P. 8 ¿Sabe quién es?		Cuando sabe quién es P. 9 ¿Sabe a qué partido pertenece? (No leer opciones)		Cuando sabe quién es P. 10 ¿Qué tan favorable o desfavorable es su opinión acerca de...? (Leer opciones)			
	Sí	No (Pasar a p. 16)	No leer opciones		Muy favorable	Favorable	Desfavorable	Muy desfavorable
			Sí identificó	No identificó				
Guadalupe Méndez Peña	1	2	(Partido 1) 1	2	4	3	2	1
Agustín Almanza Mendoza	1	2	(Partido 2) 1	2	4	3	2	1
Sandra Hernández Jiménez	1	2	(Partido 3) 1	2	4	3	2	1
Roberto Yáñez Ruiz	1	2	(Partido 4) 1	2	4	3	2	1

Preguntas abiertas

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, y puede variar de población en población.

Preguntas abiertas No delimitan las alternativas de respuesta. Son útiles cuando no hay suficiente información sobre las posibles respuestas de las personas.

EJEMPLOS

¿Por qué asiste a psicoterapia?

¿Qué opina de las medidas de apoyo a la población que adoptó el gobierno para disminuir el impacto del último terremoto ocurrido el 20 de noviembre?

¿Conviene usar preguntas cerradas o abiertas?

Un cuestionario obedece a diferentes necesidades y a un problema de investigación, lo cual origina que en cada estudio el tipo de preguntas sea distinto. Algunas veces se incluyen tan sólo preguntas cerradas, otras ocasiones 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 más fáciles de codificar y preparar para su análisis. Asimismo, estas preguntas requieren un menor esfuerzo por parte de los encuestados, que no tienen que escribir o verbalizar pensamientos, sino únicamente seleccionar la alternativa que sintetice mejor su respuesta. Responder a un cuestionario con preguntas cerradas toma menos tiempo que contestar uno con preguntas abiertas. Cuando el cuestionario se envía por correo, se tiene un mayor grado de respuesta cuando es fácil de contestar y completarlo requiere menos tiempo. Otras ventajas son: se reduce la ambigüedad de las respuestas y se favorecen las comparaciones entre las respuestas (Burnett, 2009).

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 participantes. Su redacción exige mayor laboriosidad y un profundo conocimiento del planteamiento por parte del investigador o investigadora (Vinueza, 2005).

Para *formular preguntas cerradas* es necesario anticipar las posibles alternativas de respuesta. De no ser así, es muy difícil plantearlas. Además, el investigador debe asegurarse de que los participantes a quienes 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 participantes las comprendan es muy sencillo. Pero si preguntamos sobre las razones y los motivos que provocan esa preferencia, señalar las opciones es algo más complejo.

Las *preguntas abiertas* proporcionan una información más amplia y son particularmente útiles cuando no tenemos información sobre las posibles respuestas de las personas o cuando ésta es insuficiente. También sirven en situaciones donde se desea profundizar una opinión o los motivos de un comportamiento. Su mayor desventaja es que son más difíciles de codificar, clasificar y preparar para

el análisis. Además, llegan a presentarse sesgos derivados de distintas fuentes; por ejemplo, quienes enfrentan dificultades para expresarse en forma oral y por escrito quizá no respondan con precisión a lo que en realidad desean, o generen 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; Saris y Gallhofer, 2007). Asimismo, responder a preguntas abiertas requiere de un mayor esfuerzo y de más 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 suelen ser más confiables y válidas para medir esa variable, de acuerdo con la situación del estudio (planteamiento del problema, características de la muestra, tipo de análisis a efectuar, etcétera).

Con frecuencia, las preguntas cerradas se construyen con fundamento en preguntas abiertas. Por ejemplo, en la prueba piloto puede elaborarse una pregunta abierta y posteriormente a su aplicación, sobre la base de las respuestas, se genera el ítem cerrado.



Figura 9.10 Ejemplo del paso de una pregunta abierta a la elaboración de una pregunta cerrada.

¿Una o varias preguntas para medir una variable?

En ocasiones sólo basta una pregunta para recolectar la información necesaria sobre la variable considerada. 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 se requiere elaborar varias preguntas para verificar la consistencia de las respuestas.

Por ejemplo, algunas asociaciones latinoamericanas de investigación de mercados e instituciones educativas miden el nivel socioeconómico tomando en cuenta diversas variables:⁹

⁹ Asociación Mexicana de Agencias de Investigación de Mercados y Opinión Pública, AMAI (2008), López Romo (2008) y Universidad de Celaya (2009). Para conocer el método por puntos para la ubicación del nivel socioeconómico, se recomienda consultar directamente estas fuentes o contactar a la asociación de empresas de investigación de mercados de su país.

1. Escolaridad del jefe del hogar.
2. Número de focos en la vivienda.
3. Número de habitaciones en la vivienda, sin incluir baños.
4. Número de baños con regadera.
5. Número de automóviles y otros vehículos en la cochera.
6. Posesión de ciertos aparatos (computadora, lavadora, equipo de sonido, horno de microondas, etcétera).
7. Características de la vivienda (techo estable y seguro en su vivienda —no de cartón ni hule, etc.—, piso firme en su interior —cemento, concreto o piso de mosaico—, agua que llega por/ mediante tubería a su vivienda, etcétera).

Con base en estas variables se construyen índices, cada una de ellas tiene un peso o coeficiente, y al final se otorga una puntuación que determina el nivel socioeconómico con mayor precisión. Sin embargo, esto puede resultar muy complejo para el alumno que comienza con sus primeras investigaciones, por lo cual la alternativa sería preguntar a los miembros de la familia que trabajan: ¿aproximadamente cuál es su nivel mensual de ingresos? y cuestionando: ¿cuántos focos eléctricos tiene aproximadamente en su casa?¹⁰

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 distintas dimensiones o componentes, en los cuales se incluyen preguntas para medir esas 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 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 el nivel de vida en el trabajo. Una de las variables importantes era el “grado en que se distribuía la información sobre el proceso de calidad en la organización”. Esta variable se midió a través de las siguientes preguntas:

- a) Respecto de los programas de información sobre calidad, ¿cuáles de las siguientes actividades se efectúan en esta empresa?
 1. Planeación del manejo de datos sobre calidad.
 2. Formas de control.
 3. Elaboración de reportes con datos sobre calidad.
 4. Evaluación sistemática de los datos sobre calidad.
 5. Distribución generalizada de información sobre calidad.
 6. Sistemas de autocontrol de calidad.
 7. Distribución selectiva de datos sobre calidad.
 8. Programa de comunicación interna sobre el proceso de calidad.

¹⁰ En varios estudios se ha demostrado que el nivel de ingresos está relacionado con el número de focos de una casa-habitación (residencia, hogar o propiedad). El número de focos se vincula con el número de cuartos, extensión de la casa, tamaño del jardín y otros factores, es decir, con el valor de la propiedad (Comunicometría, 1988; Universidad de Celaya, 2009). Los rangos podrían ser: 3 focos o menos: estratos muy desfavorecidos; de 4 a 5 focos: estratos desfavorecidos; de 6 a 10 focos: estratos medios típicos; 11 a 15: estratos medios favorecidos; 16 a 20: medios/altos; 21 a 30: altos favorecidos; más de 31: muy altos o completamente favorecidos. Se sugiere excluir candiles. En cada nación cambia la designación de cada estrato, y no queremos utilizar los términos “bajos”, nos parecen peyorativos.

- b) Sólo a quienes distribuyen selectivamente datos sobre calidad: ¿a qué niveles de la empresa?
 c) Sólo a quienes distribuyen selectivamente datos sobre calidad: ¿a qué funciones?
 d) ¿Qué otras actividades se realizan en esta empresa para los programas de información sobre calidad?

En este ejemplo, las preguntas *b)* y *c)* se elaboraron para ahondar 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.

Una modalidad de cuestionamientos múltiples lo es la batería de preguntas, la cual sirve para: *a)* ahorrar espacio en el cuestionario, *b)* facilitar la comprensión del mecanismo de respuesta (si se entiende de la primera pregunta, se comprenderán las demás) (Corbetta, 2003) y *c)* construir índices que permitan obtener una calificación total.

EJEMPLO

Variable a medir: visión departamental

Definición conceptual: percepción de la meta departamental en cuanto a claridad, naturaleza visionaria, grado en que es posible alcanzarla y medida en que puede ser compartida, y que representa una fuerza motivacional para el trabajo (Anderson y West, 1998; Hernández Sampieri, 2005).¹¹

Preguntas o ítems:

1. ¿Qué tan claros tiene los objetivos de su departamento?
2. ¿En qué medida considera usted que los objetivos de su departamento son útiles y apropiados?
3. ¿Qué tan de acuerdo está usted con estos objetivos de su departamento?
4. ¿En qué medida piensa usted que los objetivos de su departamento son claros?
5. ¿En qué medida piensa usted que los objetivos de su departamento son comprendidos por sus compañeros de trabajo del mismo departamento?
6. ¿En qué medida considera usted que sus compañeros de departamento están de acuerdo con los objetivos?
7. ¿En qué medida considera que los objetivos del departamento pueden lograrse actualmente?

	Completamente (mucho)	Aceptablemente	Regular	Poco	Nada
1. ¿Qué tan claros tiene los objetivos de su departamento?	5	4	3	2	1
2. ¿En qué medida considera usted que los objetivos de su departamento son útiles y apropiados?	5	4	3	2	1
3. ¿Qué tan de acuerdo está usted con estos objetivos de su departamento?	5	4	3	2	1
4. ¿En qué medida piensa usted que los objetivos de su departamento son claros?	5	4	3	2	1
5. ¿En qué medida piensa usted que los objetivos de su departamento son comprendidos por sus compañeros de trabajo del mismo departamento?	5	4	3	2	1
6. ¿En qué medida considera usted que sus compañeros de departamento están de acuerdo con los objetivos?	5	4	3	2	1
7. ¿En qué medida considera que los objetivos del departamento pueden lograrse actualmente?	5	4	3	2	1

¿Las preguntas van precodificadas o no?

Siempre que se pretenda efectuar análisis estadístico, se requiere codificar las respuestas de los participantes a las preguntas del cuestionario, y debemos recordar que esto significa asignarles símbolos o valores

¹¹ El ejemplo contiene solamente algunas de las preguntas de la escala original. Asimismo, se conjuntaron con otras escalas y se distribuyeron a lo largo del cuestionario.

numéricos y que cuando se tienen preguntas cerradas es posible codificar *a priori* o precodificar las opciones de respuesta, e incluir esta precodificación en el cuestionario (como en el último ejemplo).

EJEMPLO

De preguntas precodificadas

¿Tiene usted inversiones en la Bolsa de Valores?

1 Sí 2 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, es decir, se han precodificado. Obviamente en las preguntas abiertas la codificación se realiza después, una vez que se tienen las respuestas. Las preguntas y opciones de respuesta precodificadas poseen la ventaja de que su codificación y preparación para el análisis son más sencillas y requieren menos tiempo.

¿Qué preguntas son obligatorias?

Las preguntas llamadas demográficas o de ubicación del participante encuestado: género, edad, nivel socioeconómico, estado civil, escolaridad (nivel de estudios), religión, afiliación política, colonia, barrio o zona donde vive, pertenencia a ciertas agrupaciones, ocupación (actividad a la que se dedica), años de vivir en el lugar actual de residencia, etc. En empresas: puesto, antigüedad, área funcional donde trabaja (gerencia, departamento, dirección o equivalente), planta u oficinas donde labora, y demás preguntas. En cada investigación debemos analizar cuáles son pertinentes y nos resultarán útiles.

¿Qué 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 tienen que ser claras, precisas y comprensibles para los sujetos encuestados. Deben evitarse términos confusos, ambiguos y de doble sentido. Por ejemplo, la pregunta: “¿ve usted televisión?”, es confusa, no delimita cada cuánto. Sería mucho mejor especificar: ¿acostumbra usted ver televisión diariamente?, ¿cuántos días durante la última semana vio televisión?, y después preguntar horarios, canales y contenidos de los programas. Otro ejemplo inconveniente sería: ¿le gusta el deporte? No se sabe si se trata de verlo por televisión o en vivo, si de practicarlo o qué, y en última instancia, ¿cuál deporte? Otro caso que genera confusión son los términos con múltiples significados (Burnett, 2009), por ejemplo: ¿su empleo es estable?, implica un concepto de estabilidad de empleo que no tiene un solo significado. ¿Qué se considera estable?: ¿un contrato por un año, por dos, por cinco...?

Un caso común de confusión son las palabras sobre la temporalidad, resulta nebuloso el cuestionamiento: ¿ha asistido recientemente al cine?, ya que implica otras preguntas: ¿qué signi-

fica recientemente?, ¿ayer, la última semana, el último mes? Sería mejor interrogar: durante las últimas dos semanas (o mes), ¿cuántas veces ha ido al cine? De igual forma: ¿ha trabajado desde joven?, habrá de sustituirse por: ¿a partir de qué edad comenzó a trabajar?

- b) Es aconsejable que las preguntas sean lo más breves posible, porque las preguntas largas suelen resultar tediosas, toman más tiempo y pueden distraer al participante; pero como menciona Rojas (2001) no es recomendable sacrificar la claridad por la concisión. Cuando se trata de asuntos complicados tal vez es mejor una pregunta más larga, debido a que facilita el recuerdo, proporciona al sujeto más tiempo para pensar y favorece una respuesta más articulada (Corbetta, 2003). La directriz a seguir es que se incluyan las palabras necesarias para que se comprenda la pregunta, sin ser repetitivos o barrocos.
- c) Deben formularse con un vocabulario simple, directo y familiar para los participantes. El lenguaje debe adaptarse al habla de la población a la que van dirigidas las preguntas (Gambarra, 2002). Recuerde que es ineludible tomar en cuenta su nivel educativo y el socioeconómico, las palabras que maneja, etcétera.
- d) No pueden incomodar a la persona encuestada ni ser percibidas como amenazantes y nunca ésta debe sentir que se le enjuicia. Debemos inquirir de manera sutil. Preguntas como: ¿acostumbra consumir algún tipo de bebida alcohólica?, tienden a provocar rechazo. Es mejor cuestionar: ¿algunos de sus amigos acostumbran consumir algún tipo de bebida alcohólica?, y después utilizar preguntas tenues que indirectamente nos indiquen si la persona acostumbra consumir esta clase de bebidas (¿cuál es su tipo de bebida favorita?, ¿cada cuánto se reúne con sus amigos?, etc.). Mertens (2005) sugiere sustituir la pregunta: ¿es usted alcohólico? (en extremo amenazante), por la siguiente formulación: El consumo de bebidas como el ron, tequila, vodka y whisky en esta ciudad es de X botellas de un litro, ¿en qué medida usted estaría por encima o por debajo de esta cantidad? (alternativas de respuesta: por encima, igual o por debajo). Gochros (2005) recomienda cambiar la pregunta: ¿consume drogas?, por: ¿qué opina de las personas que consumen drogas en dosis mínimas? En estos casos de preguntas difíciles, es posible usar escalas de actitud en lugar de preguntas o aun otras formas de medición (como se verá en la parte de escalas actitudinales y en otros instrumentos). Hay temáticas en las que a pesar de que se utilicen preguntas sutiles, el encuestado se puede sentir molesto. Tal es el caso del desempleo, la homosexualidad, el SIDA, la prostitución, la pornografía, los anticonceptivos y las adicciones.
- e) Las preguntas deben referirse preferentemente a un solo aspecto o una relación lógica. Por ejemplo, la pregunta: ¿acostumbra usted ver televisión y escuchar radio diariamente?, expresa dos aspectos y llega a confundir. Es necesario dividirla en dos preguntas, una relacionada con la televisión y otra relacionada con la radio. Otro ejemplo: ¿sus padres eran saludables?, es una pregunta problemática, además del concepto “saludable” (confuso), es imposible de responder en el caso de que la madre nunca se hubiera enfermado de gravedad y nunca hubiera sido hospitalizada y, en cambio, el padre hubiera padecido severos problemas de salud.
- f) Las preguntas no habrán de inducir las respuestas. Se tienen que evitar preguntas tendenciosas o que dan pie a elegir un tipo de respuesta (directivas). Por ejemplo, ¿considera a nuestro compañero Ricardo Hernández como el mejor candidato para dirigir nuestro sindicato?, es una pregunta tendenciosa, pues induce la respuesta. Lo mismo que la pregunta: ¿los trabajadores argentinos son muy productivos? Se insinúa la respuesta en la pregunta. Resultaría mucho más conveniente interrogar: ¿qué tan productivos considera usted, en general, a los trabajadores argentinos? (y mostrar alternativas).

EJEMPLO

—¿Qué tan productivos considera usted, en general, a los trabajadores argentinos?

Sumamente
productivos

Más bien
productivos

Más bien impro-
ductivos

Sumamente
improductivos

Otros ejemplos inconvenientes serían: ¿piensa usted votar por tal partido político en las próximas elecciones?, ¿usted considera que debemos retirar las tropas de nuestro país de la coalición... para evitar amenazas a nuestra seguridad nacional? El participante nunca debe sentirse presionado. Un factor importante a considerar es la deseabilidad social, a veces las personas utilizan respuestas culturalmente aceptables. Por ejemplo, la pregunta: ¿le gustaría casarse?, podría inducir y forzar a más de una persona a responder de acuerdo con las normas de su comunidad. Resulta mejor cuestionar: ¿qué opina del matrimonio?, y más adelante inquirir sobre sus anhelos y expectativas al respecto. Una interrogante como: ¿acostumbra leer el periódico?, puede llevarnos a respuestas socialmente válidas: “sí, lo leo a diario, yo leo mucho” (cuando no es cierto). Es mejor preguntar: ¿suele tener tiempo para leer el periódico?, ¿con qué frecuencia?

- g) Las preguntas no pueden apoyarse en instituciones, ideas respaldadas socialmente ni en evidencia comprobada. Es también una manera de inducir respuestas. Por ejemplo, la pregunta: la Organización Mundial de la Salud ha realizado diversos estudios y concluyó que el tabaquismo provoca diversos daños al organismo, ¿considera usted 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...”, etc., no deben anteceder a las preguntas, ya que influyen y sesgan las respuestas.
- h) Es aconsejable evitar preguntas que nieguen el asunto que se interroga. Por ejemplo: ¿qué niveles de la estructura organizacional *no* apoyan el proceso de calidad? Es mejor preguntar sobre qué niveles *sí* apoyan el proceso. O bien: ¿qué *no* le agrada de este centro comercial?, es preferible cuestionar: ¿qué le desagrada de este centro comercial? Tampoco es conveniente incluir dobles negaciones (son positivas, pero suelen confundir): ¿considera que la mayoría de las mujeres casadas preferiría *no* trabajar si *no* tuviera presión económica? Mejor se redacta de manera positiva.
- i) No deben hacerse preguntas racistas o sexistas ni que ofendan a los participantes. Es obvio, pero no está de más recalcarlo. Se recomienda también sortear las preguntas *con fuerte carga emocional o muy complejas*, que más bien son preguntas para entrevistas cualitativas (por ejemplo: ¿cómo era la relación con su ex marido? —aunque una escala completa puede ser la solución— o, ¿qué siente usted sobre la muerte de su hijo?)
- j) En las preguntas con varias categorías de respuesta, y donde el sujeto participante sólo tiene que elegir una, llega a ocurrir que el orden en que se presentan dichas opciones afecta las respuestas de los participantes (por ejemplo, que tiendan a favorecer a la primera o a la última opción de respuesta). Entonces resulta conveniente rotar el orden de lectura de las respuestas a elegir de manera proporcional. Por ejemplo, si preguntamos: ¿cuál de los siguientes cuatro candidatos presidenciales considera usted que logrará disminuir verdaderamente la inflación? Veinticinco por ciento de las veces (o una de cada cuatro ocasiones) que se haga la pregunta se menciona primero al candidato *A*, 25% se menciona primero al candidato *B*, 25% al candidato *C* y el restante 25% al candidato *D*. Asimismo, cuando las alternativas son demasiadas es más difícil responder, por ello es conveniente limitarlas a las mínimas necesarias.

A continuación incluimos la tabla 9.4 sobre problemas al generar preguntas, adaptada de Creswell (2005).

▲ **Tabla 9.4** Ejemplos de algunos problemas al elaborar preguntas

Problema	Ejemplo de pregunta problemática	Mejora a la pregunta
Pregunta confusa por la vaguedad de los términos	¿Votará en las próximas elecciones?	Precisar términos: En las próximas elecciones del 10 de noviembre para elegir alcalde de Monterrey, ¿piensa ir a votar?
Dos o más conceptos o dos preguntas en una sola	¿Qué tan satisfecho está usted con el servicio de comedor y el servicio médico que se ofrece en la empresa?	Una pregunta por concepto: ¿Qué tan satisfecho está usted con el servicio de comedor que se ofrece en la empresa? ¿Qué tan satisfecho está usted con el servicio médico que se ofrece en la empresa?

(continúa)

▲ **Tabla 9.4** Ejemplos de algunos problemas al elaborar preguntas (*continuación*)

Problema	Ejemplo de pregunta problemática	Mejora a la pregunta
Demasiadas palabras	Como usted sabe, el próximo 10 de noviembre se celebrarán elecciones locales en este municipio de Cortázar para elegir alcalde, en esa fecha: ¿piensa usted acudir a las urnas a emitir su voto por el candidato que considera será el mejor alcalde para el municipio?	Reducir términos: En las próximas elecciones del 10 de noviembre para elegir alcalde de Cortazar, ¿piensa ir a votar?
Pregunta negativa	¿Los estudiantes no deben portar o llevar armas a la o en la escuela?	Cambiarla a neutral: ¿Los estudiantes deben o no portar armas en la escuela?
Contiene “jerga lingüística”	¿Qué tan “chida” o “padre” es la relación con su empresa?	Eliminar dicha jerga: ¿Qué tan orgulloso se encuentra usted de trabajar en esta empresa?
Se traslapan las categorías de respuesta	¿Podría indicarme su edad? __18 __ 19 __19 __ 20 __20 __ 21 __21 __ 22	Lograr que las categorías sean mutuamente excluyentes: ¿Podría indicarme su edad? __18 __ 19 __20 __ 21 __22 __ 23
Categorías de respuesta sin balance entre las favorables y las desfavorables (positivas y negativas)	¿En qué medida está usted satisfecho con su superior inmediato? <input type="checkbox"/> Insatisfecho <input type="checkbox"/> Medianamente satisfecho <input type="checkbox"/> Satisfecho <input type="checkbox"/> Sumamente satisfecho	Proporcionar equilibrio entre opciones favorables y desfavorables: ¿En qué medida está usted satisfecho con su superior inmediato? <input type="checkbox"/> Sumamente insatisfecho <input type="checkbox"/> Más bien insatisfecho <input type="checkbox"/> Ni insatisfecho ni satisfecho <input type="checkbox"/> Más bien satisfecho <input type="checkbox"/> Sumamente satisfecho
Incongruencia entre la pregunta y las opciones de respuesta	¿En qué medida está usted satisfecho con su superior inmediato? <input type="checkbox"/> Muy poco importante <input type="checkbox"/> Poco importante <input type="checkbox"/> Medianamente importante <input type="checkbox"/> Importante <input type="checkbox"/> Muy importante	Generar categorías que coincidan con la pregunta: ¿En qué medida está usted satisfecho con su superior inmediato? <input type="checkbox"/> Sumamente insatisfecho <input type="checkbox"/> Más bien insatisfecho <input type="checkbox"/> Ni insatisfecho ni satisfecho <input type="checkbox"/> Más bien satisfecho <input type="checkbox"/> Sumamente satisfecho
Sólo una parte de los participantes pueden entender la pregunta	¿Cuál es el género y marca de bebida etílica que acostumbra adquirir con un mayor índice de frecuencia en sus compras?	Simplificar términos: ¿Cuál es el tipo de bebida alcohólica y de qué marca acostumbra comprar con mayor frecuencia?
Utilización de términos en otro idioma	¿Qué efectos tuvo en esta empresa el <i>downsizing</i> ?	Traducir términos: ¿Qué efectos tuvo en esta empresa la reducción de empleados?

(continúa)

▲ **Tabla 9.4** Ejemplos de algunos problemas al elaborar preguntas (*continuación*)

Problema	Ejemplo de pregunta problemática	Mejora a la pregunta
La pregunta puede ser inadecuada para parte de la población	¿Cómo le afectó el incremento en la tasa impositiva para empleados gubernamentales?	<p>Agregar preguntas que segmenten a la población:</p> <p>¿Actualmente trabaja?</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> <p>¿Trabaja usted en...</p> <p><input type="checkbox"/> Empresa?</p> <p><input type="checkbox"/> Por cuenta propia (independiente)?</p> <p><input type="checkbox"/> Gobierno?</p> <p>Entonces, a quienes pertenezcan a la última categoría, se les pregunta:</p> <p>¿Cómo le afectó el incremento en la tasa impositiva para empleados gubernamentales?</p>

En relación con cada pregunta del cuestionario, León y Montero (2003) sugieren cuestionar: ¿es necesaria la pregunta?, ¿es lo suficientemente concreta?, ¿responderán los participantes sinceramente?

¿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 participante se adentre 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 como: ¿está de acuerdo con que se legalice el aborto en este país? Sin lugar a dudas sería un fracaso. Bostwick y Kyte (2005) y Babbie (2009) señalan que los primeros cuestionamientos deben resultar interesantes para los participantes. A veces incluso, pueden ser divertidos (por ejemplo, en la investigación de la moda y la mujer mexicana que se verá en la cuarta parte del libro sobre modelos mixtos, al comenzar a inquirir sobre los tipos de prendas que compraban las participantes, la primera pregunta fue: ¿sueles ponerte una pijama para dormir?, cuestionamiento que resultó sumamente divertido y provocó hilaridad, logrando relajar a las encuestadas. Desde luego, la pregunta la hicieron mujeres entrevistadoras jóvenes).

A veces los cuestionarios comienzan con las preguntas demográficas ya mencionadas, pero en otras ocasiones es mucho mejor hacer este tipo de preguntas al final del cuestionario, particularmente en casos donde los participantes puedan sentir que se comprometen si responden el cuestionario.

Cuando construimos un cuestionario, es indispensable que pensemos en cuáles son las preguntas ideales para iniciar. Éstas deberán lograr que el sujeto se concentre en el cuestionario. Gambará (2002) sugiere el procedimiento de “embudo” en la presentación de las preguntas: ir de las más generales a las más específicas. Una característica fundamental de un cuestionario es que las preguntas importantes nunca deben ir al final.

¿De qué está formado un cuestionario?

Además de las preguntas y categorías de respuestas, un cuestionario está formado básicamente por: portada, introducción, instrucciones insertas a lo largo del mismo y agradecimiento final.

Portada

Ésta incluye la carátula; en general, debe ser atractiva gráficamente para favorecer las respuestas. Debe incluir el nombre del cuestionario y el logotipo de la institución que lo patrocina. En ocasiones se agrega un logotipo propio del cuestionario o un símbolo que lo identifique.

Introducción

Debe incluir:

- Propósito general del estudio.
- Motivaciones para el sujeto encuestado (importancia de su participación).
- Agradecimiento.
- Tiempo aproximado de respuesta (un promedio o rango). Lo suficientemente abierto para no presionar al participante, pero tranquilizarlo.
- Espacio para que firme o indique su consentimiento (a veces se incluye al final o en ocasiones es innecesario).
- Identificación de quién o quiénes lo aplican.
- Explicar brevemente cómo se procesarán los cuestionarios y una cláusula de confidencialidad del manejo de la información individual.
- Instrucciones iniciales claras y sencillas (cómo responder en general, con ejemplos si se requiere).

Cuando el cuestionario se aplica mediante entrevista, la mayoría de tales elementos son explicados por el entrevistador. El cuestionario debe ser y parecer corto, fácil y atractivo.

A continuación, se presenta un ejemplo de carta introductoria y otro de instrucciones generales para responder al cuestionario.

EJEMPLO

Carta introductoria

Buenos días (tardes):

Estamos trabajando en un estudio que servirá para elaborar una tesis profesional acerca de la biblioteca de la Universidad de Celaya.

Quisiéramos pedir tu ayuda para que contestes algunas preguntas que no llevarán mucho tiempo. Tus respuestas serán confidenciales y anónimas. No hay preguntas delicadas.

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 se puede responder a una opción; otras son de varias opciones y también se incluyen preguntas abiertas.

Muchas gracias por tu colaboración.

EJEMPLO

Instrucciones de un cuestionario ENCUESTA DEL CLIMA ORGANIZACIONAL

INSTRUCCIONES

Emplee un lápiz o un bolígrafo de tinta negra para rellenar el cuestionario. Al hacerlo, piense en lo que sucede la mayoría de las veces en su trabajo.

No hay respuestas correctas o incorrectas. Éstas simplemente reflejan su opinión personal.

Todas las preguntas tienen cinco opciones de respuesta, elija la que mejor describa lo que piensa usted. Solamente una opción.

Recordemos que en ocasiones se presentan tarjetas con las opciones de respuestas y se instruye al entrevistador para que las muestre a los participantes. Por ejemplo:

—Hablando de la mayoría de sus proveedores, en qué medida conoce usted... (*Mostrar la tarjeta uno y marcar la respuesta en cada caso.*)

EJEMPLO

De pregunta con tarjeta de respuestas

	Completamente (5)	Bastante (4)	Regular (3)	Poco (2)	Nada (1)
• ¿Las políticas de su proveedor?					
• ¿Sus finanzas (estados financieros)?					
• ¿Los objetivos de su área de ventas?					
• ¿Sus programas de capacitación para vendedores?					
• ¿Número de empleados de su área de ventas?					
• ¿Problemas laborales?					
• ¿Los métodos de producción que tienen?					
• ¿Otros clientes de ellos?					
• ¿Su índice de rotación de personal?					

Las instrucciones son tan importantes como las preguntas y es necesario que sean claras para los usuarios a quienes van dirigidas.

Agradecimiento final

Aunque haya agradecido de antemano, vuelva a agradecer la participación.

Formato, distribución de instrucciones, preguntas y categorías

Las preguntas deben estar organizadas para que sea más fácil de responder el cuestionario. Es importante asegurarnos de numerar páginas y preguntas.

La manera en que pueden distribuirse preguntas, categorías de respuesta e instrucciones es variada. Algunos prefieren colocar las preguntas a la izquierda y las respuestas a la derecha, con lo que se tendría un formato como el siguiente:

EJEMPLO**Modelo de formato de distribución de preguntas**

- ¿Considera a su jefe o superior inmediato como su amigo? Definitivamente sí Sí No Definitivamente no
- ¿Cuando tiene problemas se siente apoyado por su jefe o superior inmediato? Definitivamente sí Sí No Definitivamente no
- ¿Considera que su jefe o superior inmediato le orienta adecuadamente en su trabajo? Definitivamente sí Sí No Definitivamente no
- ¿Tiene una buena impresión de su jefe o superior inmediato? Definitivamente sí Sí No Definitivamente no

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

EJEMPLO**Modelo de formato horizontal**

PRESENTACIÓN

Preguntas sobre el superior inmediato

¿Considera a su jefe o superior inmediato como su amigo?

Definitivamente sí Sí No Definitivamente no

¿Cuándo tiene problemas se siente apoyado por su jefe o superior inmediato?

Definitivamente sí Sí No Definitivamente no

¿Considera que su jefe o superior inmediato le orienta adecuadamente en su trabajo?

Definitivamente sí Sí No Definitivamente no

¿Tiene una buena impresión de su jefe o superior inmediato?

Definitivamente sí Sí No Definitivamente no

PREGUNTAS SOBRE MOTIVACIÓN

Otros combinan diversas posibilidades, distribuyendo preguntas que miden la misma variable a través de todo el cuestionario. Cada quien es capaz de utilizar el formato que desee o juzgue conveniente, lo importante es que en su totalidad sea comprensible para el usuario: que las instrucciones, preguntas y respuestas se diferencien; que el formato no resulte visualmente tedioso y se lea sin dificultad.

Hoy en día resulta común elaborar cuestionarios en CD y otros medios como Palm, PC (contestar directamente en aparatos portátiles y cuestionarios electrónicos, etc.), así como diseños de éstos para páginas web y blogs en internet que contienen fotografías, dibujos, secuencias de video y música. Son sumamente atractivos.

¿De qué tamaño debe ser un cuestionario?

No existe una regla al respecto, pero si es muy corto se pierde información y si resulta largo llega a ser tedioso. En este último caso, las personas se negarían a responder o, al menos, lo contestarían en forma incompleta. La abuela doña Margarita Castelán Sampieri repetía el refrán: “lo bueno y breve, doblemente bueno”. El tamaño depende del número de variables y dimensiones a medir, el interés de los participantes y la manera como se administre (de este punto se hablará en el siguiente apartado). Cuestionarios que duran más de 35 minutos suelen resultar fatigosos, a menos que los sujetos estén muy motivados para contestar (por ejemplo, cuestionarios de personalidad o cuestionarios para obtener un trabajo). Una recomendación que ayuda a evitar un cuestionario más largo de lo requerido es: *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 participantes a los cuales se les aplicaron, o al menos las principales tendencias de respuestas en una muestra de los cuestionarios aplicados. Es importante anotar que esta actividad es similar a “cerrar” una pregunta abierta por medio de la prueba piloto, pero el producto es diferente. En este caso, con la codificación de preguntas abiertas se obtienen ciertas categorías que representan los resultados finales.

El procedimiento consiste en encontrar y dar nombre a los patrones generales de respuesta (respuestas similares o comunes), listar estos patrones y después asignar un valor numérico o un símbolo a cada patrón. Así, un patrón constituirá una categoría de respuesta. Para cerrar las preguntas abiertas se sugiere el siguiente procedimiento:

1. Seleccionar determinado número de cuestionarios mediante un método adecuado de muestreo, que asegure la representatividad de los participantes investigados.
2. Observar la frecuencia con que aparece cada respuesta a determinadas preguntas.
3. Elegir las respuestas que se presentan con mayor frecuencia (patrones generales de respuesta).
4. Clasificar las respuestas elegidas en temas, aspectos o rubros, de acuerdo con un criterio lógico, cuidando que sean mutuamente excluyentes.
5. Darle un nombre o título a cada tema, aspecto o rubro (patrón general de respuesta).
6. Asignarle el código a cada patrón general de respuesta.

Por ejemplo, en la investigación de *Comunicometría* (1988) se hizo una pregunta abierta: ¿de qué manera la alta gerencia busca obtener la cooperación del personal para el desarrollo del proyecto de calidad?

Las respuestas fueron múltiples, pero se encontraron los patrones generales de respuesta que se muestran en el ejemplo.

EJEMPLO

De codificación de preguntas abiertas

Códigos	Categorías (patrones o respuestas con mayor frecuencia de mención)	Frecuencia de mención
1	Involucrando al personal y comunicándose con él	28
2	Motivación e integración	20
3	Capacitación en general	12
4	Incentivos/recompensas	11
5	Difundiendo el valor “calidad” o la filosofía de la empresa	7
6	Grupos o sesiones de trabajo	5

Códigos	Categorías (patrones o respuestas con mayor frecuencia de mención)	Frecuencia de mención
7	Posicionamiento del área de calidad o equivalente	3
8	Sensibilización en grupo	2
9	Desarrollo de la calidad de vida en el trabajo	2
10	Incluir aspectos de calidad en el manual de inducción	2
11	Poner énfasis en el cuidado de la maquinaria	2
12	Trabajar bajo un buen clima laboral	2
13	Capacitación “en cascada”	2
14	Otras	24

Como varias categorías o diversos patrones tenían únicamente dos frecuencias, éstos a su vez se redujeron a sólo seis, como se ejemplifica a continuación.

EJEMPLO

De reducción o agrupamiento de categorías

Códigos	Categorías (frecuencias)
1	Involucrando al personal y comunicándose con él (28)
2	Motivación e integración/mejoramiento del clima laboral (22)
3	Capacitación (14)
4	Incentivos/recompensas (11)
5	Difundiendo el valor “calidad” o la filosofía de la empresa (7)
6	Grupos o sesiones de trabajo (7)
7	Otras (33)

Al “cerrar” preguntas abiertas y codificarlas, debe tenerse en cuenta que un mismo patrón de respuesta puede expresarse con diferentes palabras. Por ejemplo, ante la pregunta ¿qué sugerencias haría para mejorar al programa *Estelar*? Las respuestas: mejorar las canciones y la música, cambiar las canciones, incluir nuevas y mejores canciones, etc., se agruparían en la categoría o el patrón de respuesta *modificar la musicalización del programa*.

¿En qué contextos puede administrarse o aplicarse un cuestionario?

Los cuestionarios se aplican de dos maneras fundamentales: autoadministrado y por entrevista (personal o telefónica).

1. Autoadministrado

Autoadministrado significa que el cuestionario se proporciona directamente a los participantes, quienes lo contestan. No hay intermediarios y las respuestas las marcan ellos. Pero la forma de autoadministra-

ción puede tener distintos contextos: *individual, grupal o por envío (correo tradicional, correo electrónico y página web o blog)*.

En el caso individual, el cuestionario se entrega al participante y éste lo responde, ya sea que acuda a un lugar para hacerlo (como ocurre cuando se llena un formulario para solicitar empleo) o lo conteste en su lugar de trabajo, hogar o estudio. Por ejemplo, si los participantes fueran una muestra de directivos de laboratorios farmacéuticos de Bogotá, se acudiría a sus oficinas y se les entregarían los cuestionarios. Los ejecutivos se autoadministrarían el cuestionario y esperaríamos a que lo respondan (caso poco común) o lo recolectaríamos otro día. El reto de esta última situación es lograr que los participantes devuelvan el cuestionario contestado completamente. Es conveniente que quien lo entregue posea habilidades para relacionarse con las personas, sea asertivo, y además se caracterice por una elevada persistencia. En nuestra experiencia, en distintos países de Iberoamérica, jóvenes de ambos géneros con buena capacidad comunicativa logran porcentajes de recuperación por encima de 90% en tiempos aceptables (una semana o menos). Y no es necesario que sean físicamente atractivos (aunque ayuda), más bien el éxito reside en su motivación y tenacidad. Asimismo, el mayor coste o gasto de esta clase de administración de los cuestionarios lo representa su distribución y recolección.

En el segundo caso, se reúne a los participantes en grupos (a veces pequeños —cuatro a seis personas—, otras en grupos intermedios —entre siete y 20 sujetos—, incluso en grupos grandes de 21 a 40 individuos). Por ejemplo, empleados (en encuestas de clima organizacional es muy común juntar a grupos de 25, entregarles el cuestionario, introducir al propósito del estudio y al instrumento, responder dudas y pedirles que al concluir lo depositen en una urna sellada, para mantener la confidencialidad), padres de familia (en reuniones escolares), televidentes (cuando asisten a un foro televisivo), alumnos (en sus salones de clase), etc. Es tal vez la forma más económica de aplicar un cuestionario.

A continuación se incluye en la tabla 9.6 una lista de verificación de los aspectos centrales para administrar cuestionarios en grupo.

▲ **Tabla 9.6** Listado de puntos a verificar al administrar cuestionarios en grupo¹³

1. ¿Tenemos suficientes cuestionarios?	Sí _____	No _____
2. ¿Hemos diseñado alguna medida para que quienes no puedan asistir a la sesión respondan al cuestionario?	Sí _____ ¿Cuál? _____	No _____
3. ¿Se notificó formalmente a los participantes potenciales la fecha, la hora y el lugar en que se aplicaría el cuestionario?	Sí _____ ¿Cómo? (carta, correo electrónico, memorándum)	No _____
4. ¿Se verificó que el lugar donde se aplicará el cuestionario presenta las condiciones adecuadas de espacio e iluminación?	Sí _____ ¿Quién verificó? _____	No _____
5. ¿Se tomaron acciones para aislar el lugar de fuentes potenciales de ruido u otras distracciones?	Sí _____	No _____

(continúa)

¹³ Adaptado de McMurthy (2005).

▲ **Tabla 9.6** Listado de puntos a verificar al administrar cuestionarios en grupo (*continuación*)

6. ¿Quién va a leer en voz alta las instrucciones y asistir a los participantes a lo largo de la sesión?	Sí _____ Persona(s): _____	No _____
7. ¿Las instrucciones incluyen cómo responder al cuestionario?	Sí _____	No _____
8. ¿Se contempló un tiempo razonable para responder dudas e inquietudes de los participantes antes de que comiencen a contestar el cuestionario?	Sí _____	No _____
9. ¿Quien va a leer en voz alta las instrucciones tiene una voz nítida y suficientemente fuerte para que todos le escuchen y su lectura será pausada?	Sí _____ Persona(s): _____	No _____
10. ¿Se verificará que todos hayan respondido al cuestionario?	Sí _____	No _____
11. ¿Quién dará las gracias a los participantes por su cooperación?	Sí _____ Persona(s): _____	No _____
12. ¿Quién enviará las cartas de agradecimiento o equivalentes a los participantes y a quienes facilitaron la sesión?	Sí _____ Persona(s): _____	No _____

Los cuestionarios para autoadministración deben ser particularmente atractivos (a colores, en papel especial, con diseño original, etc.; si el presupuesto lo permite).

En el caso de autoadministración por envío, se manda el cuestionario a los participantes por correo postal privado o mensajería (por la rapidez), por medio del correo electrónico, también se les puede pedir que ingresen a una página web o blog para responderlo.

Por correo tradicional: postal o servicio de paquetería o mensajería especializada. El cuestionario se envía junto con una carta explicativa firmada por el investigador o investigadores, la cual hace las funciones de la introducción (con los elementos comentados previamente: propósito del estudio, motivadores, agradecimiento, tiempo de respuesta, etc., excepto las instrucciones que suelen incluirse en el instrumento). Si la carta va membretada con el logotipo del instrumento, mejor.

Se recomienda que los cuestionarios sean más cortos. Si al hablar de otros instrumentos autoadministrados, se comentó que las instrucciones deben ser precisas y claras, esto resulta particularmente importante en estos casos, ya que las posibilidades de realimentación y resolución de dudas se reducen al mínimo. La carátula, además de lo que se señaló previamente, debe contener la fecha exacta de envío. En las instrucciones es necesario agregar la fecha en que se requiere sea devuelto y la forma de regresar el cuestionario contestado, paso a paso. De ser posible, resulta aconsejable designar a una persona para que atienda dudas y comentarios del instrumento y el estudio, por medio telefónico y/o correo electrónico, obviamente tienen que proporcionar sus datos completos. Ofrecer a los participantes un resumen de los resultados, una vez que concluya la investigación, es una práctica recomendable (la cual se puede enviar por correo electrónico).

Asimismo, el paquete enviado a cada individuo potencial incluye dos sobres: uno que contiene el cuestionario y la carta, y el otro para que devuelva el cuestionario cumplimentado. Desde luego, este último con los datos completos del remitente (destinatario final) y con el porte de regreso o la guía de paquetería prepagada (necesitamos cubrir todos los gastos generados en este proceso). Un diseño original de los sobres puede ser de gran ayuda, al menos para que sean abiertos.

Es fundamental contactar vía telefónica y/o correo electrónico al futuro encuestado, para motivarlo a que conteste el cuestionario. Una vez que se reciba su respuesta, es preciso agradecerle su cooperación. Algunas personas se niegan a participar en investigaciones, porque fueron tratadas con descortesía una vez que se obtuvo de ellas lo que se deseaba.

Los cuestionarios autoadministrados pueden ser procesados de forma casi inmediata si se usa codificación por lectura óptica. Es decir, si el papel del cuestionario cubre ciertos requisitos y es respondido con un lápiz o bolígrafo especial. Se ahorra una la codificación, puesto que el sistema lee las respuestas y automáticamente las envía a la base de datos correspondiente.

Por correo electrónico. Se trata de un procedimiento similar, lo único que cambia es el medio de contacto. La carta, carátula, instrucciones y el cuestionario son enviados a través de un correo electrónico (*e-mail*).

Por medio de una página de internet. Esta vía es similar, en cuanto a la mecánica, a las dos anteriores. Pero en este caso se le pide al participante (por contacto telefónico o correo electrónico) que acceda a un sitio web, donde se localiza el cuestionario, el cual se contesta en el momento o por etapas; otra modalidad puede ser que se “descargue” o “baje” el cuestionario para guardarlo como archivo en la computadora y posteriormente, una vez contestado, se envía por correo electrónico.

Los cuestionarios utilizados en medios electrónicos regularmente se elaboran en un programa de texto e imagen, o se escanean (si están impresos con anterioridad) y “se anexan” en el correo electrónico (como un “archivo adjunto”), también se pueden colocar o “subir” al sitio web, aunque para este segundo caso lo más común es que se elaboren especialmente para tal ambiente. En ambas situaciones, las posibilidades de diseño del instrumento son amplísimas.

Las limitaciones de los estudios que utilizan el correo electrónico y la web, residen en que no todas las personas poseen computadora e internet (sobre todo en América Latina) y algunos individuos (por ejemplo, los mayores de 60 años) se resisten a utilizar estos recursos, porque es una tecnología reciente y desconocida para gran parte de ellos.

Una tasa de devolución de cuestionarios cumplimentados por correo o de manera electrónica por encima de 50% es muy favorable (Mertens, 2005).

Una posibilidad novedosa son las entrevistas interactivas (algunas son modalidades telefónicas, otras se trata de los denominados “medios inteligentes” de correo electrónico o de sitios web), en donde un sistema se contacta vía telefónica o por correo electrónico con los participantes potenciales y efectúa la administración del cuestionario o lo envía. Son mecanismos con reconocimiento de voz, lectura óptica y dictado digital. El problema es —hasta el momento— que la mayoría de la gente se da cuenta de que no es otro ser humano con quien entra en contacto y suele negarse a responder. Además, la saturación de correos electrónicos, llamadas telefónicas y sitios web hacen difícil captar la atención de los participantes potenciales. Si se utilizan, se aconseja que los cuestionarios sean muy breves, no más de 10 preguntas. Desde luego esta situación irá modificándose y cada vez serán más los estudios que utilicen tales tecnologías.

Por otro lado, los sitios web que presentan encuestas de opinión rápida, donde las personas accesan a páginas en las que pueden responder el cuestionario, tienen serios problemas de muestreo (desde luego, se trata de muestras no probabilísticas), esto se debe a que, como ya dijimos, no toda la población puede hacerlo, con lo cual quedan excluidos diversos segmentos, al igual que personas sumamente ocupadas o que simplemente no se interesan en contestar.

En este sentido, Cook, Heath y Thompson (2000 y 2001) realizaron un par de estudios que se centraron en la utilización de internet, cuyos resultados se aplican a todas las vertientes de autoadministración de cuestionarios por envío. De este modo, resultó que tres factores son clave para obtener elevados índices de retorno de cuestionarios: *a)* seguimiento persistente a casos de no respuesta,

b) vinculación de manera personalizada con los participantes y c) contacto antes del envío. La tasa de retorno es mayor en cuestionarios cortos que en los largos.

Una ventaja de estos métodos es que cuando se hacen preguntas personales o de mayor carga emotiva, el sujeto puede contestar de manera más relajada y sincera, puesto que no está frente a otra persona. Vinuesa (2005) señala que la encuesta por correo permite una selección muestral de los participantes de acuerdo con su perfil sociodemográfico, de compra, estilo de vida, etc., y de individuos concretos (profesionales, miembros de alguna asociación, etcétera).

Algunas desventajas residen en que nunca podremos estar seguros de quién respondió el cuestionario y la ausencia de un encuestador impide asegurar la franqueza de las respuestas.

Es importante no realizar investigaciones que requieran enviar el cuestionario en épocas complejas del año (vacaciones de verano o invierno: en Navidad la saturación es impresionante) o que lo sean para la población en estudio (por ejemplo, a fiscalistas y contadores de empresas durante momentos de cierres contables y pago de impuestos; a las personas de edad avanzada en épocas de frío extremo, etcétera).

Para ahondar en el tema de las aplicaciones de cuestionarios por internet y correo, recomendamos a Dillman, Smyth y Christian (2009).

2. Por entrevista personal

Las *entrevistas* implican que una persona calificada (entrevistador) aplica el cuestionario a los participantes; el primero hace las preguntas a cada entrevistado y anota las respuestas. Su papel es crucial, es una especie de filtro.

El primer contexto que revisaremos de una entrevista es el personal (“cara a cara”).

Normalmente se tienen varios entrevistadores, quienes deberán estar capacitados en el arte de entrevistar y conocer a fondo el cuestionario. Quienes no deberán sesgar o influir en las respuestas, por ejemplo, reservarse de expresar aprobación o desaprobación respecto de las respuestas del entrevistado, reaccionar de manera ecuánime cuando los participantes se perturben, contestar con gestos ambiguos cuando los sujetos busquen generar una reacción en ellos, etc. Su propósito es lograr que se culmine exitosamente cada entrevista, evitando que decaiga la concentración e interés del participante, además de orientar a éste en el tránsito del instrumento. Las explicaciones que proporcione deberán ser breves pero suficientes. Tiene que ser neutral, pero cordial y servicial. Asimismo, es muy importante que transmita a todos los participantes que no hay respuestas correctas o equivocadas. Por otra parte, su proceder debe ser lo más estándar posible (mismos señalamientos, presentación uniforme, etc.). Con respecto a las instrucciones del cuestionario, algunas son para el entrevistado y otras para el entrevistador. Este último debe recordar que al inicio se comenta: el propósito general del estudio, las motivaciones y el tiempo aproximado de respuesta, agradeciendo de antemano la colaboración.

Estamos de acuerdo con León y Montero (2003), quienes manifiestan que el anterior método descrito es el que consigue un mayor porcentaje de respuestas a las preguntas, su estimación es de 80 a 85%. Incluso puede ser superior a esta cifra con una planeación adecuada.

En relación con el perfil de entrevistadores no hay un consenso, por ejemplo, Corbetta (2003) sugiere que sean mujeres casadas, amas de casa, de mediana edad, diplomadas y de clase media. León y Montero (2003) recomiendan que sean siempre profesionales. En nuestra experiencia el tipo de entrevistador depende del tipo de persona entrevistada. Por ejemplo, que pertenezca a un nivel socioeconómico similar a la mayoría de la muestra, sea joven y haya cursado asignaturas o materias de investigación, que posea facilidad de palabra y capacidad de socializar. Como ya se explicitó previamente, los estudiantes de ambos géneros funcionan mucho mejor. Por tanto, es claro que para este fin deben rechazarse personas inseguras o excesivamente tímidas.

Rogers y Bouey (2005), así como Moule y Goodman (2009), diferencian entre la entrevista cuantitativa y la cualitativa; en relación con la primera, mencionan las siguientes características:

- a) El principio y final de la entrevista se definen con claridad. De hecho, tal definición se integra en el cuestionario.

- b)* El mismo instrumento es aplicado a todos los participantes, en condiciones lo más similares que sea posible.
- c)* El entrevistador pregunta, el entrevistado responde.
- d)* Se busca que sea individual, sin la intrusión de otras personas que pueden opinar o alterar de alguna manera la entrevista.
- e)* Es poco a nada anecdótica (aunque en algunos casos es recomendable que el entrevistador anote cuestiones fuera de lo común como ciertas reacciones y negativas a responder).
- f)* La mayoría de las preguntas suelen ser cerradas, con mínimos elementos rebatibles, ampliaciones y sondeos.
- g)* El entrevistador y el propio cuestionario controlan el ritmo y la dirección de la entrevista.
- h)* El contexto social no es un elemento a considerar, lo es solamente el ambiental.
- i)* El entrevistador procura que el patrón de comunicación sea similar (su lenguaje, instrucciones, etcétera).

Desde luego, se trata de entrevistas cuya naturaleza es muy distinta y a veces opuesta. Sin embargo, recomendamos que se complemente la lectura de estas líneas con la de entrevistas cualitativas en el capítulo 14 “Recolección y análisis de los datos cualitativos”.

Asimismo, la capacitación de entrevistadores debe incluir cuestiones de comunicación no verbal básicas (control de gestos, manejo de silencios, etc.), además de todos los puntos que se revisaron anteriormente.

Cabe señalar que, cuando se trata de *entrevista personal*, el lugar donde se realice es importante (oficina, hogar o casa-habitación, sitio público, como centro comercial, parque, escuela, etc.). Por ejemplo, Jaffé, Pasternak y Grifel (1983) hicieron 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, aunque la entrevista en los puntos de compraventa es menos costosa. En cualquier caso se aconseja que se busque un lugar lo más discreto, silencioso y privado que sea posible. A la misma conclusión llegaron Hernández Sampieri, Cuevas y Méndez (2009), quienes entre 2007 y 2009 hicieron ocho encuestas para conocer la intención del voto y las tendencias electorales en varios municipios de México, y encontraron resultados similares al entrevistar en el hogar y en sitios públicos (parques, mercados, centros comerciales, etcétera).

En estas entrevistas es común mostrar visualmente las opciones de respuesta a los entrevistados, mediante tarjetas, en especial cuando se incluyen más de cinco o son complejas. Pongamos de ejemplo la siguiente tarjeta.¹⁴

EJEMPLO

De tarjeta para mostrar al entrevistado cuando hay diversas opciones de respuesta ¿Cuáles considera usted que son los tres principales problemas en este municipio?

- | | | |
|--|---|--|
| <input type="checkbox"/> Pandillerismo | <input type="checkbox"/> Desempleo | <input type="checkbox"/> Inseguridad en las calles |
| <input type="checkbox"/> Venta de drogas-narcomenudeo | <input type="checkbox"/> Falta de vivienda | <input type="checkbox"/> Problemas en la recolección de basura |
| <input type="checkbox"/> Pobreza | <input type="checkbox"/> Falta de infraestructura (calles, puentes, etcétera) | <input type="checkbox"/> Escasez de agua |
| <input type="checkbox"/> Corrupción de funcionarios de la alcaldía | <input type="checkbox"/> Empleo mal remunerado | <input type="checkbox"/> Carencia de servicios de salud |

¹⁴ El ejemplo se ha simplificado por cuestiones de espacio, las opciones fueron obtenidas después de una prueba piloto, se trata de una encuesta hecha en un municipio de Colombia.

También hace algunos años se generó un sistema para sustituir al cuestionario (de lápiz y papel), que es el CAPI (Computer-Assisted Personal Interviewing), en donde el entrevistador muestra al participante una computadora u ordenador personal portátil (*notebook* o *laptop*) que contiene el cuestionario y este último responde guiado por el primero. A veces, la computadora tiene forma de un pequeño pizarrón plano y no posee teclado (de 20 a 40 centímetros de largo y alto), entonces se le presenta al entrevistado el instrumento (a colores, con video, imágenes y muchas más posibilidades) y lo contesta utilizando una pluma electrónica.

Casi siempre las entrevistas son individuales, aunque podrían aplicarse a un grupo pequeño (si ésta fuera la unidad de análisis o caso). Es decir, el cuestionario lo responden entre todos sus miembros o parte de ellos (por ejemplo, cuestionarios para parejas o una familia, o un departamento o de una empresa).

Los cuestionarios aplicados por CAPI, Palm y otros dispositivos similares tienen la ventaja de que los datos se capturan y agregan a la base de datos de forma automática, de manera que en cualquier momento podemos hacer un corte y efectuar toda clase de análisis (vea tendencias, evaluar funcionamiento del instrumento, etc.) (Hernández Sampieri y Mendoza, 2008). La desventaja es obvia: el costo; el cual es muy difícil de absorber por parte de un estudiante o maestro, e incluso, una institución.¹⁵

Idealmente, después de una entrevista se puede preparar un informe que indique: si el participante se mostraba sincero, la manera como respondió, el tiempo que duró la entrevista, el lugar donde se realizó, las características del entrevistado, los contratiempos que se presentaron y la forma en que se desarrolló la entrevista, así como otros aspectos que se consideren relevantes.

3. Por entrevista telefónica

Obviamente, la diferencia con el anterior tipo de entrevista es el medio de comunicación, que en este caso es el teléfono (hogar, oficina, móvil o celular). Las entrevistas telefónicas son la forma más rápida de realizar una encuesta. Junto con la aplicación grupal de cuestionarios es la manera más económica de aplicar un instrumento de medición, con la posibilidad de asistir a los sujetos de la muestra. Ha sido muy utilizada en los países desarrollados debido a la vertiginosa evolución de la telefonía.

Las habilidades requeridas de parte de los entrevistadores son parecidas a las de la entrevista personal, excepto que éstos no tienen que confrontarse “cara a cara” con los participantes (no importa la vestimenta ni el aspecto físico, pero sí la voz, su modulación y claridad son fundamentales). El nivel de rechazo suele ser menor que la entrevista frente al participante, con excepción de periodos de “saturación telefónica”. Por ejemplo, cuando las compañías de un ramo compiten en cuestiones de mercadeo telefónico; tal como ocurrió en varios países latinoamericanos con la apertura comercial a nuevas empresas telefónicas (dichos consorcios iniciaron una campaña para contactar a todos los números telefónicos del país a cualquier hora con el fin de ofrecer sus servicios, llamando a los hogares desde los domingos a las siete de la mañana o después de las 10 de la noche entre semana y hasta en la madrugada). Otro caso es el periodo de elecciones en países donde no se legisla el telemercado, los equipos de los candidatos contendientes llegan a aturdir a los ciudadanos mediante comunicaciones telefónicas en busca del voto y para efectuar encuestas de tendencias.

Una ventaja enorme de este método reside en que se puede acceder a barrios inseguros, a conjuntos exclusivos y edificios o casas donde se impide el ingreso (León y Montero, 2003), así como a lugares geográficamente lejanos al investigador.

Algunas recomendaciones para las entrevistas telefónicas son las siguientes:¹⁶

1. Enviar una carta, telefonema o correo electrónico previo, en que se indique el objetivo de la entrevista, la persona o institución que realiza la encuesta y el día y hora en que se efectuará la comunicación telefónica.

¹⁵ Aunque en la Universidad de Celaya están investigando la forma de bajar costos y hacer el sistema accesible, incluso con financiamiento público. Tal vez pronto se tengan noticias vía la página de este libro.

¹⁶ Mertens (2005); Hernández Sampieri y Mendoza (2008); y Kuusela, Callegaro y Vehovar (2009).

2. Realizar la comunicación en el día y hora acordada.
3. El entrevistador debe identificarse y recordarle al entrevistado el propósito del estudio; asimismo, debe asegurarse que es un buen momento para la comunicación.
4. Entre la comunicación previa y la entrevista telefónica no debe pasar más de una semana (programar llamadas adecuadamente).
5. El entrevistador debe asegurarse que está hablando con la persona correcta o que posee el perfil adecuado según la definición de la muestra.
6. Indicar el tiempo que tomará la entrevista.
7. Utilizar un cuestionario breve con preguntas preferentemente estructuradas (cerradas) y sencillas. Más de 15 a 17 preguntas suelen complicar la situación.
8. El entrevistador debe vocalizar correctamente y a la misma velocidad de su interlocutor.
9. Anotar casos de rechazos y las razones.
10. En el entrenamiento, simular las condiciones de aplicación (igual en la prueba piloto).
11. Establecer metas de comunicaciones telefónicas por hora.
12. Si se pretende grabar la entrevista, debe interrogarse al participante si está de acuerdo.

Desde luego, estas recomendaciones aplican a una encuesta telefónica donde poseemos tiempos amplios para llevarla a cabo. Pero en ocasiones, tenemos que realizar sondeos inmediatos para obtener tendencias en la opinión pública y algunas de estas recomendaciones no son pertinentes. Por ejemplo, después de una catástrofe (como un magnicidio, acto terrorista o desastre natural), una noticia mundial (la elección de un nuevo Papa, un acuerdo de paz) o local (un triunfo electoral, un nuevo impuesto). Tal es el caso de las encuestas que se realizaron en los días posteriores al asesinato de J.F. Kennedy (Sheatsley y Feldman, 1964), las efectuadas después de los ataques terroristas del 11 de septiembre de 2001 en Nueva York (University of Southern California y Bendixen & Associates, 2002), las hechas con motivo de los atentados ocurridos en Madrid el 11 de marzo de 2004 (Michavila, 2005) o las ulteriores a las explosiones del 7 de julio de 2005 en Londres (COMPAS, 2005; The Harris Poll, 2005; British Broadcasting Corporation, 2005), así como la encuesta telefónica nacional llevada a cabo en México después del brote del virus de la influenza humana (Consulta Mitofsky, 2009).

Para la administración de encuestas telefónicas se dispone de varias tecnologías, además de las ya comentadas de reconocimiento de voz y dictado digital, como el CATI (Computer-Assisted Telephone Interviewing), en donde el entrevistador se sienta frente a su computadora u ordenador, cuyo sistema selecciona números telefónicos generados al azar y los marca automáticamente. Una vez que contesta la persona indicada, él comienza a leer las preguntas en el monitor y anota las respuestas (desde luego, mediante el teclado o el *mouse*), las cuales son capturadas y codificadas de manera automática. El sistema gestiona el desarrollo de la entrevista, ya que va remitiendo a las opciones adecuadas (en caso de preguntas condicionadas, como por ejemplo: ¿tiene usted una cuenta en este banco, sí o no? Si la respuesta es “sí”, entonces continúa con la siguiente pregunta enlazada: ¿qué servicios usa...? Pero si la respuesta fue un “no”, puede concluir con un muchas gracias..., o pasar a otras preguntas). El entrevistador puede utilizar diademas con audífonos y micrófono. O bien, el sistema tiene la facilidad de reconocimiento de voz y de capturar directamente la respuesta. Es una interfase con el *Random Digit Dialing*.

Una enorme desventaja de las encuestas telefónicas es que están limitadas a una cuantas preguntas o no se pueden efectuar mediciones complejas de variables o profundizar en ciertos temas. Pero una vez más, los datos se capturan y agregan a la base de datos de forma automática y se pueden hacer cortes de la información de manera inmediata y realizar toda clase de análisis.

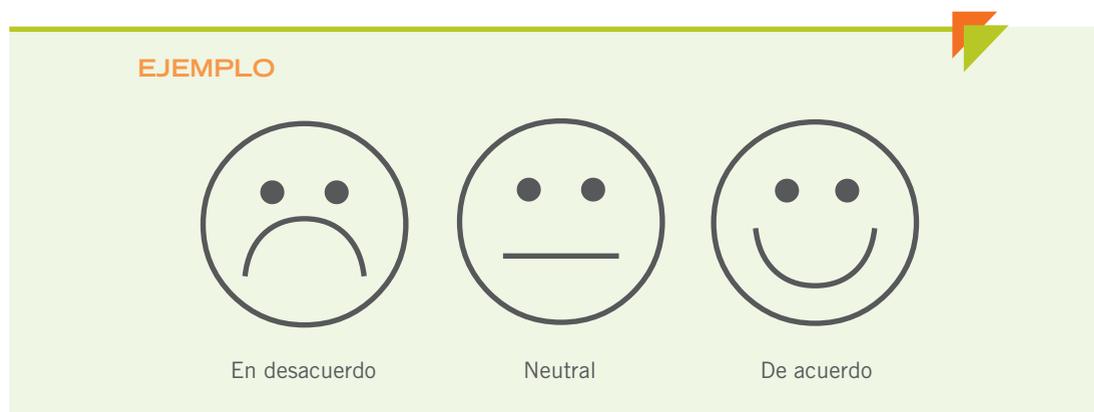
Corbetta (2003) sugiere que si la pregunta se va a presentar oralmente (mediante entrevista) las preguntas no contengan más de cinco opciones de respuesta, ya que por encima de este límite se suelen olvidar las primeras.

Cuando se realizan entrevistas personales en el hogar o telefónicas se debe tomar en cuenta el *horario*. Ya que si efectuamos la visita o hablamos por teléfono sólo a una hora (digamos en la mañana), nos encontraremos con unos cuantos subgrupos de la población (por ejemplo, amas de casa).

Una variación de la administración de cuestionarios por teléfono es la siguiente: en un programa radial o televisivo se solicita la opinión o respuesta de los receptores a una pregunta o algunas cuantas preguntas, éstos deben marcar un número telefónico y contestar las opciones de respuesta con las que concuerden más. El problema de estas encuestas reside en la muestra, que desde luego no es probabilística, sino que se conforma de voluntarios que cubren dos condiciones: tener teléfono y estar viendo o escuchando la emisión del programa. Este proceder nos conduce más que a un estudio, a un sondeo. Lo cual no es un error en sí, lo grave es que se pretenda generalizar los resultados a una población (por ejemplo, los habitantes de una ciudad, un estado, provincia o departamento; o peor aún, un país).

Algunas consideraciones adicionales para la administración del cuestionario

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 para las opciones de respuestas. Como en el siguiente ejemplo.



Con trabajadores de niveles de lectura básica se recomienda utilizar entrevistas o cuestionarios autoadministrados sencillos que se apliquen en grupos, con la asesoría de entrevistadores o supervisores capacitados.

En algunos casos, con ejecutivos que difícilmente vayan a dedicarle a un solo asunto más de 20 minutos, se pueden utilizar cuestionarios autoadministrados o entrevistas telefónicas. Con estudiantes suelen funcionar los cuestionarios autoadministrados.

Algunas asociaciones realizan encuestas por correo y ciertas empresas envían cuestionarios a sus ejecutivos y supervisores mediante el servicio interno de mensajería o por correo electrónico. Cuando el cuestionario contiene unas cuantas preguntas (su administración toma entre cuatro y cinco 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 los cuestionarios se están aplicando correctamente.

La elección del contexto para administrar el cuestionario deberá ser muy cuidadosa y dependerá del presupuesto disponible, el tiempo de entrega de los resultados, el planteamiento del problema, la naturaleza de los datos y el tipo de participantes (edad, nivel educativo, etcétera).

A continuación incluimos la tabla 9.7 que compara de manera sencilla las formas de administración.

▲ **Tabla 9.7** Comparación de las principales formas de administración de cuestionarios

Método de administración	Tasa de respuesta	Presupuesto o coste (fuente que origina el mayor gasto)	Rapidez con que se administra	Profundidad de los datos obtenidos	Tamaño del cuestionario
Autoadministrado (individual)	Media	Medio (pago de recolectores)	Media	Alta	Cualquier tamaño razonable
Autoadministrado (grupal)	Alta	Bajo (sesiones)	Rápido	Alta	Cualquier tamaño razonable
Autoadministrado (envío correo o paquetería)	Baja	Bajo por correo postal(envíos) Medio por paquetería (envíos)	Lenta	Alta	Cualquier tamaño razonable
Autoadministrado por correo electrónico o página web	Baja	Bajo (diseño electrónico)	Media	Alta	Cualquier tamaño razonable
Entrevista personal	Alta	Elevado (pago a entrevistadores y gastos de viaje)	Media	Alta	Cualquier tamaño razonable
Entrevista telefónica	Alta	Bajo (llamadas telefónicas locales y entrevistadores)	Rápido	Baja	Corto

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

OO4 Escalas para medir las actitudes

Actitud Predisposición aprendida para responder coherentemente de manera favorable o desfavorable ante un objeto, ser vivo, actividad, concepto, persona o sus símbolos.

Una **actitud** es una predisposición aprendida para responder coherentemente de una manera favorable o desfavorable ante un objeto, ser vivo, actividad, concepto, persona o sus símbolos (Fishbein y Ajzen, 1975; Haddock y Maio, 2007; y Oskamp y Schultz, 2009). Así, los seres humanos tenemos actitudes hacia muy diversos objetos, símbolos, etc.; por ejemplo, actitudes hacia el aborto, la política económica, la familia, un profesor, diferentes grupos étnicos, la ley, nuestro trabajo, una nación específica, los osos, 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, 2000). Si detecto que la actitud de un grupo hacia la contaminación es desfavorable, esto no significa que las personas estén tomando acciones para evitar contaminar el ambiente, aunque sí es un indicador de que pueden adoptarlas en forma paulatina. La actitud es como una “semilla” que bajo ciertas condiciones suele “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. A continuación examinaremos los primeros dos, que son los utilizados con mayor frecuencia. En el capítulo 7 del CD anexo: “Recolección de los datos cuantitativos, segunda parte”, se comenta el tercer método: escalograma de Guttman.



Escalamiento tipo Likert

Este método fue desarrollado por Rensis Likert en 1932; sin embargo, se trata de un enfoque vigente y bastante popularizado.¹⁷ Consiste en un conjunto de ítems presentados en forma de afirmaciones o juicios, ante los cuales se pide la reacción de los participantes. Es decir, se presenta cada afirmación y se solicita al sujeto que externé su reacción eligiendo uno de los cinco puntos o categorías de la escala. A cada punto se le asigna un valor numérico. Así, el participante obtiene una puntuación respecto de la afirmación y al final su puntuación total, sumando las puntuaciones obtenidas en relación con todas las afirmaciones.

Las afirmaciones califican al objeto de actitud que se está midiendo. El objeto de actitud puede ser cualquier “cosa física” (un vestido, un automóvil...), un individuo (el Presidente, un líder histórico, mi madre, mi sobrino Alexis, un candidato a una elección...), un concepto o símbolo (patria, sexualidad, la mujer vallenata —Colombia—, el trabajo), una marca (Adidas, Ford...), una actividad (comer, beber café...), una profesión, un edificio, etc. Por ejemplo, Kafer *et al.* (1989) generaron varias escalas para medir las actitudes hacia los animales y Meerkerk *et al.* (2009) desarrollaron un instrumento basado en escalas Likert para determinar la severidad del uso compulsivo de internet.

Tales frases o juicios deben expresar sólo una relación lógica; además, es muy recomendable que no excedan de 20 palabras.

Escalamiento Likert Conjunto de ítems que se presentan en forma de afirmaciones para medir la reacción del sujeto en tres, cinco o siete categorías.

EJEMPLO

De frase

Objeto de actitud medido

El voto

Afirmación

“Votar es una obligación de todo ciudadano responsable”

En este caso, la afirmación incluye ocho palabras y expresa una sola relación lógica (X–Y). Las opciones de respuesta o puntos de la escala son cinco e indican cuánto se está de acuerdo con la frase correspondiente.¹⁸ Las opciones más comunes se presentan en la figura 9.11. Debe recordarse que a cada una de ellas se le asigna un valor numérico (precodificado o no) y sólo puede marcarse una respuesta. Se considera un dato inválido si se marcan dos o más opciones.

Las opciones de respuesta o categorías pueden colocarse de manera horizontal, como en la figura 9.11, o en forma vertical.

- () Muy de acuerdo
- () De acuerdo
- () Ni de acuerdo ni en desacuerdo
- () En desacuerdo
- () Muy en desacuerdo

¹⁷ Para conocer los orígenes de esta técnica se recomienda consultar a Likert (1976a o 1976b), Seiler y Hough (1976) y particularmente el libro original: Likert (1932).

¹⁸ Likert (1932), Futrell *et al.* (1998), Clark (2000) y Roberts y Jowell (2008).



Figura 9.11 Opciones o puntos en las escalas Likert.

O bien, utilizando recuadros en lugar de paréntesis:

- Definitivamente sí
- Probablemente sí
- Indeciso
- Probablemente no
- Definitivamente no

Es indispensable señalar que el número de categorías de respuesta debe ser igual para todas las afirmaciones. Pero siempre respetando el mismo orden o jerarquía de presentación de las opciones para todas las frases (ver tabla 9.8).

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; de este modo, cuanto más de acuerdo con la frase estén los participantes, su actitud será igualmente más favorable.

▲ **Tabla 9.8** Opciones jerárquicamente correctas e incorrectas en un ejemplo¹⁹

Objeto de actitud: mi novia	
Correcto	Incorrecto (no se respeta la misma jerarquía en todos los ítems)
<p>“Me gusta estar mucho con mi novia”</p> <p><input type="checkbox"/> Definitivamente sí</p> <p><input type="checkbox"/> Probablemente sí</p> <p><input type="checkbox"/> Indeciso</p> <p><input type="checkbox"/> Probablemente no</p> <p><input type="checkbox"/> Definitivamente no</p>	<p>“Me gusta estar mucho con mi novia”</p> <p><input type="checkbox"/> Probablemente sí</p> <p><input type="checkbox"/> Indeciso</p> <p><input type="checkbox"/> Definitivamente sí</p> <p><input type="checkbox"/> Probablemente no</p> <p><input type="checkbox"/> Definitivamente no</p>
<p>“Si por mí fuera, todos los días estaría con mi novia”</p> <p><input type="checkbox"/> Definitivamente sí</p> <p><input type="checkbox"/> Probablemente sí</p> <p><input type="checkbox"/> Indeciso</p> <p><input type="checkbox"/> Probablemente no</p> <p><input type="checkbox"/> Definitivamente no</p>	<p>“Si por mí fuera, todos los días estaría con mi novia”</p> <p><input type="checkbox"/> Definitivamente sí</p> <p><input type="checkbox"/> Probablemente sí</p> <p><input type="checkbox"/> Probablemente no</p> <p><input type="checkbox"/> Definitivamente no</p> <p><input type="checkbox"/> Indeciso</p>
<p>“Amo demasiado a mi novia”</p> <p><input type="checkbox"/> Definitivamente sí</p> <p><input type="checkbox"/> Probablemente sí</p> <p><input type="checkbox"/> Indeciso</p> <p><input type="checkbox"/> Probablemente no</p> <p><input type="checkbox"/> Definitivamente no</p>	<p>“Amo demasiado a mi novia”</p> <p><input type="checkbox"/> Definitivamente sí</p> <p><input type="checkbox"/> Probablemente sí</p> <p><input type="checkbox"/> Indeciso</p> <p><input type="checkbox"/> Probablemente no</p> <p><input type="checkbox"/> Definitivamente no</p>

EJEMPLO

“El Ministerio de Hacienda ayuda al contribuyente a resolver sus problemas en el pago de impuestos”.

En este ejemplo, si estamos “muy de acuerdo” con la afirmación implica una actitud más favorable hacia el Ministerio de Hacienda que si estamos solamente “de acuerdo”. En cambio, si estamos “muy en desacuerdo” implica una actitud muy desfavorable. Por 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, en este ejemplo, estar más de acuerdo implica una puntuación mayor.

Pero si la afirmación es *negativa*, significa que califica desfavorablemente al objeto de actitud, y cuanto más de acuerdo estén los participantes con la frase, implica que su actitud es menos favorable, esto es, más desfavorable.

¹⁹ Es un ejemplo simple para ilustrar el concepto.

EJEMPLO

“El Ministerio de Hacienda se caracteriza por obstaculizar al contribuyente en el pago de impuestos”.

En este nuevo ejemplo, si estamos “muy de acuerdo” implica una actitud más desfavorable que si estamos de “acuerdo”, y así en forma sucesiva. 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”. Cuanto más de acuerdo estemos con el juicio, nuestra actitud hacia Luis será menos favorable. Es decir, estar más de acuerdo implica una puntuación menor. *Cuando las afirmaciones son negativas se califican al contrario de las positivas.*

- (1) Totalmente de acuerdo
- (2) De acuerdo
- (3) Ni de acuerdo ni en desacuerdo
- (4) En desacuerdo
- (5) Totalmente en desacuerdo

En la figura 9.12 se presenta un ejemplo de una escala Likert para medir la actitud hacia un organismo tributario.²⁰

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, qué 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 | 4. En desacuerdo |
| 2. De acuerdo | 5. Muy en desacuerdo |
| 3. Ni de acuerdo ni en desacuerdo | |
| 2. La Dirección General de Impuestos Nacionales se caracteriza por la deshonestidad de sus funcionarios. | |
| 1. Muy de acuerdo | 4. En desacuerdo |
| 2. De acuerdo | 5. Muy en desacuerdo |
| 3. Ni de acuerdo ni en desacuerdo | |
| 3. Los servicios que presta la Dirección General de Impuestos Nacionales en general son muy buenos. | |
| 5. Muy de acuerdo | 2. En desacuerdo |
| 4. De acuerdo | 1. Muy en desacuerdo |
| 3. Ni de acuerdo ni en desacuerdo | |
| 4. La Dirección General de Impuestos Nacionales informa claramente sobre cómo, dónde y cuándo pagar los impuestos. | |
| 5. Muy de acuerdo | 2. En desacuerdo |
| 4. De acuerdo | 1. Muy en desacuerdo |
| 3. Ni de acuerdo ni en desacuerdo | |

Figura 9.12 Muestra de una escala Likert.

²⁰ El ejemplo fue utilizado en un país latinoamericano y su confiabilidad total fue de 0.89; aquí se presenta una versión reducida de la escala original. El nombre del organismo tributario que aquí se utiliza es ficticio.

5. La Dirección General de Impuestos Nacionales es muy lenta en la devolución de impuestos pagados en exceso.	1. Muy de acuerdo	4. En desacuerdo
	2. De acuerdo	5. Muy en desacuerdo
	3. Ni de acuerdo, ni en desacuerdo	
6. La Dirección General de Impuestos Nacionales informa oportunamente sobre cómo, dónde y cuándo pagar los impuestos.	5. Muy de acuerdo	2. En desacuerdo
	4. De acuerdo	1. Muy en desacuerdo
	3. Ni de acuerdo ni en desacuerdo	
7. La Dirección General de Impuestos Nacionales tiene normas y procedimientos bien definidos para el pago de impuestos.	5. Muy de acuerdo	2. En desacuerdo
	4. De acuerdo	1. Muy en desacuerdo
	3. Ni de acuerdo ni en desacuerdo	
8. La Dirección General de Impuestos Nacionales tiene malas relaciones con la gente porque cobra impuestos muy altos.	1. Muy de acuerdo	4. En desacuerdo
	2. De acuerdo	5. Muy en desacuerdo
	3. Ni de acuerdo ni en desacuerdo	

Figura 9.12 Muestra de una escala Likert (*continuación*).

Como puede observarse en la figura 9.12, 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 alcanzados respecto de cada frase. Por ello se denomina *escala aditiva*. La figura 9.13, la cual se basa en la figura 9.12, constituiría un ejemplo de cómo calificar una escala Likert.

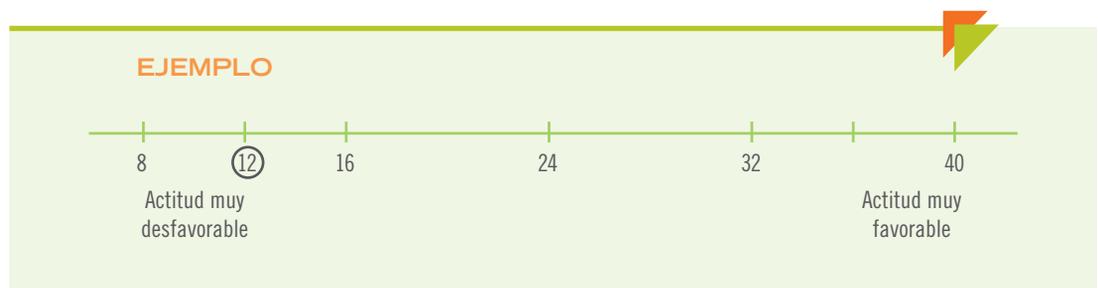
1. El personal de la Dirección General de Impuestos Nacionales es grosero al atender al público.	✗ Muy de acuerdo	4. En desacuerdo
	2. De acuerdo	5. Muy en desacuerdo
	3. Ni de acuerdo ni en desacuerdo	
2. La Dirección General de Impuestos Nacionales se caracteriza por la deshonestidad de sus funcionarios.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

Figura 9.13 Muestra de puntuaciones de la escala Likert.

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 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 <input checked="" type="checkbox"/> 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	<input checked="" type="checkbox"/> En desacuerdo 1. 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 2. De acuerdo 3. Ni de acuerdo ni en desacuerdo	4. En desacuerdo 5. Muy en desacuerdo
Valor = 1 + 2 + 1 + 3 + 1 + 1 + 2 + 1 = 12	

Figura 9.13 Muestra de puntuaciones de la escala Likert (*continuación*).

Una puntuación se considera alta o baja según el número de ítems o afirmaciones. Por ejemplo, en la escala para evaluar la actitud hacia el organismo tributario, la puntuación mínima posible es de ocho ($1 + 1 + 1 + 1 + 1 + 1 + 1 + 1$) y la máxima es de 40 ($5 + 5 + 5 + 5 + 5 + 5 + 5 + 5$), porque hay ocho afirmaciones. La persona del ejemplo obtuvo 12. Su actitud hacia el organismo tributario es más bien bastante desfavorable; veámoslo gráficamente:



Si alguien hubiera tenido una puntuación de 37 ($5 + 5 + 4 + 5 + 5 + 4 + 4 + 5$) su actitud se calificaría como sumamente favorable. En las *escalas Likert* a veces se califica el promedio resultante 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 ($12/8 = 1.5$).

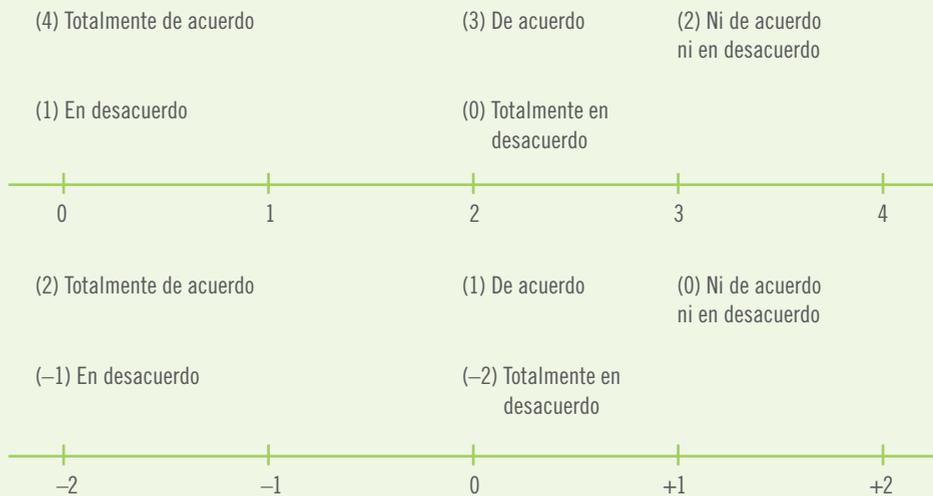
EJEMPLO



La escala Likert es, en sentido estricto, una medición ordinal; sin embargo, es común que se le trabaje como si fuera de intervalo. Creswell (2005) y Pell (2005) señalan que debe considerarse en un nivel de medición por intervalos porque ha sido probada en múltiples ocasiones. Pero otros autores, como Jamieson (2004), consideran que tiene que concebirse como ordinal y analizarse como tal. Para profundizar en esta polémica recomendamos a Hodge y Gillespie (2003), así como a Carifio y Rocco (2007 y 2008) y Achyar (2008).

Asimismo, a veces se utiliza un intervalo de 0 a 4 o de -2 a +2, en lugar de 1 a 5. Pero esto no importa porque se cambia el marco de referencia de la interpretación. Veámoslo gráficamente.

EJEMPLO



Simplemente se ajusta el marco de referencia; pero el rango se mantiene y las categorías continúan siendo cinco.

Otras condiciones sobre la escala Likert

A veces se disminuye o se incrementa el número de categorías, sobre todo cuando los participantes potenciales tienen una capacidad muy limitada de discriminación o, por el contrario, muy amplia.

EJEMPLOS

1 De acuerdo	0 En desacuerdo	
3 De acuerdo	2 Ni de acuerdo ni en desacuerdo	1 En desacuerdo
7 Totalmente de acuerdo	6 De acuerdo	5 Indeciso, pero más bien de acuerdo
4 Indeciso, ni de acuerdo ni en desacuerdo	3 Indeciso pero más bien en desacuerdo	
2 En desacuerdo	1 Totalmente en desacuerdo	

Si los participantes tienen poca capacidad de discriminar se pueden considerar dos o tres categorías. Por el contrario, si son personas con un nivel educativo elevado y gran capacidad de discriminación, pueden incluirse siete o más categorías. Pero debe recalcarse que el número de categorías de respuesta tiene que ser el mismo para todos los ítems. Si son tres, son tres categorías para todos los ítems o las afirmaciones. Si son cinco, son cinco categorías para todos los reactivos. En ocasiones se elimina la opción o categoría intermedia y neutral (ni de acuerdo ni en desacuerdo, neutral, indeciso...) para comprometer al sujeto o forzarlo a que se pronuncie de manera favorable o desfavorable.

Asimismo, como señalan Hodge y Gillespie (2003), algunos respondientes gradúan su intensidad en un continuo que va del “fuertemente de acuerdo” a “neutral” y hasta el “fuertemente en desacuerdo”, mientras que otros entienden a esta categoría central como un “no sé” o “no aplica”. Estos individuos ven al punto neutral o medio como una extensión de la dimensión de contenido, considerándolo como una opción de respuesta cuando no poseen suficiente información. En este caso, resulta apropiado ignorar tales respuestas cuando se calcula la puntuación total (Raaijmakers *et al.*, 2000; Hodge y Gillespie, 2003). Si después de la prueba piloto se observa que una cuarta parte o más de los respondientes tienden a irse a la categoría neutral en un ítem, es necesario revisarlo e incluso eliminarlo. Si esto ocurre en varios reactivos, se debe eliminar tal categoría o revisar a fondo la escala.

Un aspecto muy importante de la escala Likert es que asume que los ítems o las afirmaciones miden la actitud hacia un único concepto subyacente. En el caso de que se midan actitudes hacia varios objetos, deberá incluirse una escala por objeto, porque aunque se presenten conjuntamente, 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 con un elevado número de afirmaciones que califiquen al objeto de actitud y se administran a un grupo piloto para obtener las puntuaciones del grupo en cada ítem o frase. Estas puntuaciones se correlacionan con las del grupo a toda la escala (la suma de las puntuaciones de todas las afirmaciones), y las frases o reactivos, cuyas puntuaciones se correlacionen significativamente con las puntuaciones de toda la escala, se seleccionan para integrar el instrumento de medición. Asimismo, debe calcularse la confiabilidad y validez de la escala.

Preguntas en lugar de afirmaciones

En la actualidad, la escala original con frases se ha extendido a preguntas y observaciones. Como se puede observar en el siguiente ejemplo para evaluar al conductor de un programa televisivo.

EJEMPLO**¿Cómo considera usted al conductor del programa...?**

- 5 Muy buen conductor 4 Buen conductor 3 Regular
 2 Mal conductor 1 Muy mal conductor

Otro ejemplo sería un conjunto de preguntas formuladas en una investigación para analizar la relación de compraventa en empresas de la Ciudad de México (Paniagua, 1985). De ella se presenta un fragmento en la tabla 9.9.²¹

▲ **Tabla 9.9** Ejemplo de la escala Likert aplicada a varias preguntas

¿Para elegir a sus proveedores qué tan importante es...	Indispensable (5)	Sumamente importante (4)	Medianamente importante (3)	Poco importante (2)	No se toma en cuenta (1)
el precio?	5	4	3	2	1
la forma de pago (contado/ crédito)?	5	4	3	2	1
el tiempo de entrega?	5	4	3	2	1
el lugar de entrega?	5	4	3	2	1
la garantía del producto?	5	4	3	2	1
el prestigio del producto?	5	4	3	2	1
el prestigio de la empresa proveedora?	5	4	3	2	1
el cumplimiento del proveedor con las especificaciones?	5	4	3	2	1
la información que sobre el producto proporcione el proveedor?	5	4	3	2	1
el tiempo de trabajar con el proveedor?	5	4	3	2	1
la entrega del producto en las condiciones acordadas?	5	4	3	2	1
la calidad del producto?	5	4	3	2	1

Las respuestas se califican del modo que ya hemos comentado.

La escala en la pregunta

En ocasiones la escala se incluye en la pregunta. Mertens (2005) las denomina *preguntas actitudinales*, por ejemplo: ¿está usted fuertemente a favor, más bien a favor, más bien en contra o fuertemente en contra del aborto cuando la mujer ha sido violada?

²¹ Estas preguntas se han seguido utilizando en estudios más actuales; siguen siendo vigentes.

En la pregunta se eliminó la categoría central o intermedia. Pero estas interrogantes suelen limitarse a entrevistas de unas cuantas preguntas, porque requieren cierta capacidad de memorización.

Método de completar las frases

Hodge y Gillespie (2003) desarrollaron una derivación del escalamiento clásico de Likert, en la cual se incluyen frases incompletas respecto al objeto de actitud y a éstas se les agrega un continuo que sirve como base para las respuestas claves. Estos autores plantearon un continuo con 11 puntos o categorías (0 a 10) y que se “ancla” en cada extremo con terminaciones respecto a la frase a la que hacen referencia, las cuales representan la ausencia del constructo (cero) y la máxima “cantidad” o “presencia” del mismo (10). Sostienen que los participantes usan un número para guiarse en sus respuestas, y la frase introductoria los orienta en el continuo. Se les pide que circunscriban o marquen el número que refleje mejor su respuesta. El constructo se mide por ítems que enfatizan la fuerza del atributo. Los números trabajan en concordancia con las frases para implicar el grado de presencia de éste. El ejemplo, lo sería la *actitud intrínseca hacia la religión*.²²

EJEMPLO

De una escala donde se completan las afirmaciones

Actitud intrínseca hacia la religión

1. Mis creencias religiosas afectan:

Ningún aspecto de
mi vida

Absolutamente todos los
aspectos de mi vida

0	1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	---	----

2. Estoy consciente de la presencia de Dios...

Continuamente

Nunca

10	9	8	7	6	5	4	3	2	1	0
----	---	---	---	---	---	---	---	---	---	---

3. Respecto a las preguntas que tengo sobre la vida, mi religión responde...

Absolutamente a ninguna
de mis preguntas

Absolutamente a todas
mis preguntas

0	1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	---	----

4. Mi religión es...

El motivo más
importante de mi vida,
dirige todos los demás
aspectos

No es un factor en
mi vida

10	9	8	7	6	5	4	3	2	1	0
----	---	---	---	---	---	---	---	---	---	---

5. Leo libros, busco información en internet y veo programas que se relacionen con mi fe...

Nunca

Todos los días,
sin fallar

0	1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	---	----

²² Adaptado al español y modificado después de una prueba piloto (Hodge y Gillespie, 2003, p. 52).

6. Busco momentos para meditar y pensar sobre mi religión y Dios...

Todos los días,
sin fallar

Nunca

10	9	8	7	6	5	4	3	2	1	0
----	---	---	---	---	---	---	---	---	---	---

Al construir una escala Likert, debemos asegurarnos de que las afirmaciones y alternativas de respuesta serán comprendidas por los participantes a los que se les aplicará y que éstos tendrán la capacidad de discriminación requerida. Ello se evalúa cuidadosamente en la prueba piloto. Las escalas pueden ser autoadministradas o aplicadas mediante entrevistas, en este último caso, es recomendable mostrar al entrevistado una tarjeta donde se presenten las alternativas de respuestas o categorías. Asimismo, las escalas Likert pueden integrarse dentro de un cuestionario.

Diferencial semántico

El **diferencial semántico** fue desarrollado originalmente por Osgood, Suci y Tannenbaum (1957) para explorar las dimensiones del significado.²³ Consiste en una serie de adjetivos extremos que califican al objeto de actitud, ante los cuales se solicita la reacción del participante. Es decir, éste debe calificar al objeto de actitud a partir de un conjunto de adjetivos bipolares; entre cada par de éstos, se presentan varias opciones y la persona selecciona aquella que en mayor medida refleje su actitud.

Diferencial semántico Serie de pares de adjetivos extremos que sirven para calificar al objeto de actitud, ante los cuales se pide la reacción del sujeto, al ubicarlo en una categoría por cada par.

EJEMPLO

Escala bipolar

Objeto de actitud: Candidato "A"

justo : ____ : ____ : ____ : ____ : ____ : ____ : injusto

Debe observarse que los adjetivos son "extremos" y que entre ellos hay siete opciones de respuesta. Cada participante 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:

justo: ____ : ____ : ____ : ____ : ____ : ____ : X : injusto

²³ Para profundizar en el diferencial semántico se recomienda consultar: Osgood, Suci y Tannenbaum (1957, 1976a y 1976b), así como Heise (1976).

Si el participante 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):

justo: _____ : **X** : _____ : _____ : _____ : _____ : _____ : injusto

justo: _____ : _____ : _____ : _____ : _____ : **X** : _____ : injusto

Si el interviniente considera que el objeto de actitud se relaciona *medianamente* 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, cuanto más injusto lo considero más me acerco al extremo opuesto.

Algunos casos de adjetivos bipolares se muestran en el siguiente ejemplo. Desde luego hay muchos más que se han utilizado o que pudieran pensarse. La elección de adjetivos depende del objeto de actitud a calificar, ya que se requiere que los adjetivos se puedan aplicar a éste.

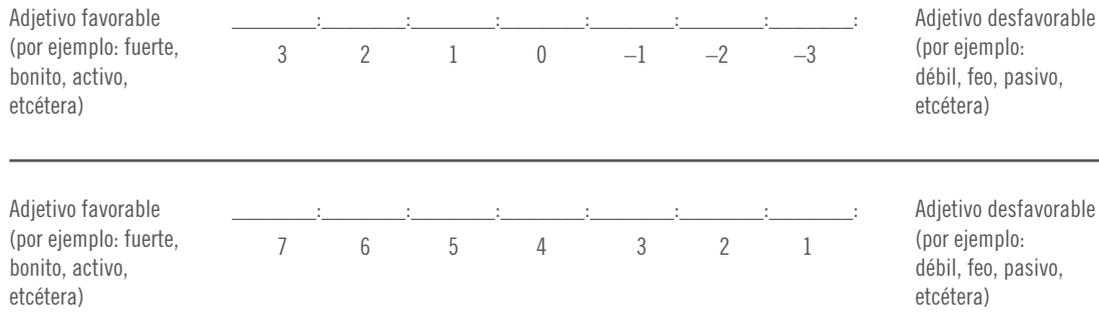
EJEMPLOS

Adjetivos bipolares

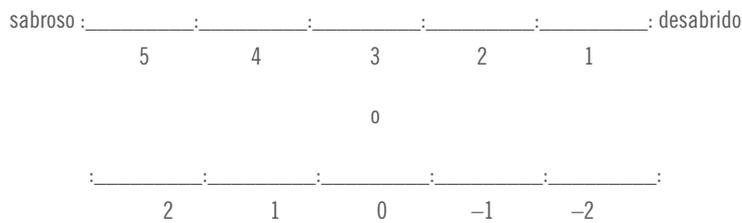
fuerte-débil	poderoso-impotente
grande-pequeño	vivo-muerto
bonito-feo	joven-viejo
alto-bajo	rápido-lento
claro-oscuro	gigante-enano
caliente-frío	perfecto-imperfecto
costoso-barato	agradable-desagradable
activo-pasivo	bendito-maldito
seguro-peligroso	arriba-abajo
bueno-malo	útil-inútil
dulce-amargo	favorable-desfavorable
profundo-superficial	asertivo-tímido
agresivo-pacífico	honesto-deshonesto
sincero-hipócrita	bien intencionado-mal intencionado

Codificación de las escalas

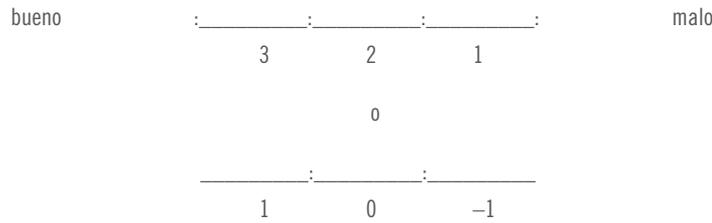
Los puntos o las categorías de la escala pueden codificarse de diversos modos, que se presentan en la figura 9.14.



En los casos en que los respondientes tengan menor capacidad de discriminación, se pueden reducir las categorías a cinco opciones. Por ejemplo:



o aun a tres opciones (lo cual es poco común):



También pueden agregarse calificativos a los puntos o las categorías de la escala (Babbie, 1979, p. 411).

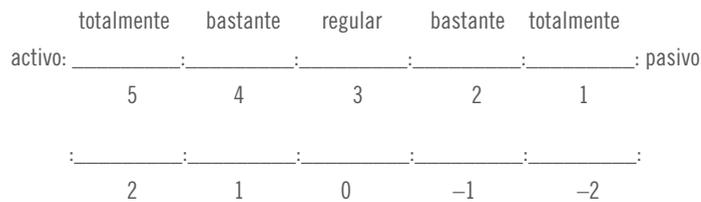
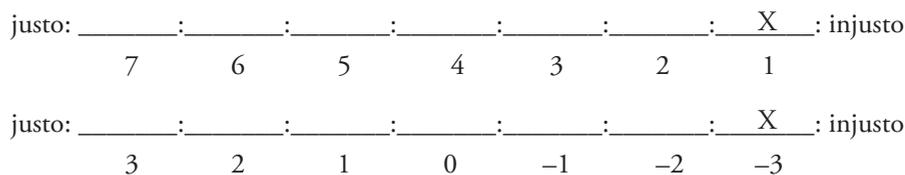
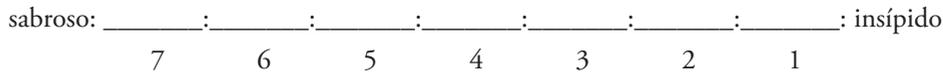


Figura 9.14 Maneras comunes de codificar el diferencial semántico.

Codificar de 1 a 7 o de -3 a 3 no tiene importancia, siempre y cuando estemos conscientes del marco de interpretación. Por ejemplo, si una persona califica al objeto de actitud: candidato “A” en la escala justo-injusto, marcando la categoría más cercana al extremo “injusto”, la puntuación sería “1” o “-3”.



Por ejemplo:



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 utilizados en investigaciones similares a la nuestra (contextos parecidos).
2. *Construimos una versión preliminar* de la escala y la administramos a un grupo de participantes a manera de prueba piloto.
3. *Correlacionamos las respuestas de los intervinientes para cada par de adjetivos o ítems.* Así, correlacionamos un ítem con todos los demás (cada par de adjetivos frente al resto).
4. *Calculamos la confiabilidad y la validez de la escala total* (todos los pares de adjetivos).
5. *Seleccionamos los ítems que presenten correlaciones significativas y discriminen entre casos* con los demás ítems. Naturalmente, si hay confiabilidad y validez, estas correlaciones serán significativas.
6. *Desarrollamos la versión final de la escala.*

La escala definitiva se califica de igual manera que la de Likert: sumando las puntuaciones obtenidas respecto de cada ítem o par de adjetivos. La figura 9.16 es un ejemplo de ello.

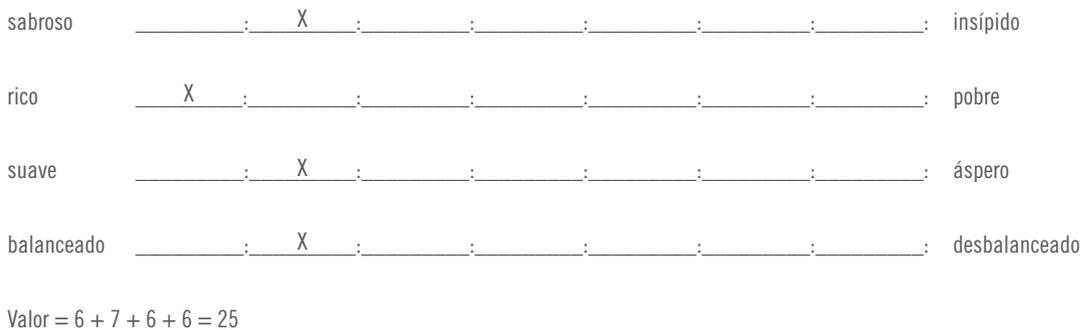


Figura 9.16 Ejemplo de cómo calificar un diferencial semántico.

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

$$\left(\frac{\text{puntuación total}}{\text{número de ítems}} \right)$$

Se pueden utilizar distintas escalas o diferenciales semánticos para medir actitudes hacia varios objetos. Por ejemplo, es posible medir con cuatro pares de adjetivos la actitud hacia el candidato “A”, con otros tres pares de adjetivos la actitud respecto de 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 con tres diferentes conceptos (“objetos de actitud”).

El *diferencial semántico* es una escala de medición ordinal, aunque es común que se le trabaje como si fuera de intervalo (Key, 1997), por las mismas razones de Likert.

Escalograma de Guttman

Escalograma de Guttman Técnica para medir las actitudes que al igual que Likert se fundamenta en juicios, ante los cuales los participantes deben externar su opinión seleccionando uno de los puntos o categorías de la escala respectiva.

El escalograma de Guttman es otra técnica para medir las actitudes y al igual que Likert se fundamenta en afirmaciones o juicios respecto del concepto u objeto de actitud, ante los cuales los participantes deben externar su opinión seleccionando uno de los puntos o categorías de la escala respectiva. Una vez más, a cada categoría se le otorga un valor numérico. Así, el participante obtiene una puntuación respecto de la afirmación y al final su puntuación total, sumando las puntuaciones obtenidas en relación con todas las afirmaciones.

La diferencia con el método de Likert es que las frases tienen diferentes intensidades (se escalan por tal intensidad), por ejemplo, la siguiente afirmación en relación con el aborto (actitud evaluada): “si en estos momentos me embarazara, jamás abortaría”; es más intensa que esta otra: “si una de mis mejores amigas se embarazara, nunca le recomendaría abortar”, y a su vez, esta última resulta más intensa que la afirmación: “si una compañera del salón de clases se embarazara, probablemente no le recomendaría abortar”. Es decir, se basa en el principio de que algunos ítems indican en mayor medida la fuerza o intensidad de la actitud.



Por cuestiones de espacio, el escalograma de Guttman no se comenta en la parte impresa de este libro, sino en el CD anexo: Material complementario → Capítulos → Capítulo 7: “Recolección de los datos cuantitativos: segunda parte”.

Otros métodos cuantitativos de recolección de los datos

¿Qué otras maneras existen para recolectar los datos desde la perspectiva del proceso cuantitativo?

QA4



En la investigación disponemos de otros métodos para recolectar los datos, tan útiles y fructíferos como los cuestionarios y las escalas de actitudes, los cuales solamente se enuncian en este capítulo, pero se comentan con mayor profundidad en el CD anexo: Material complementario → Capítulos → Capítulo 7: “Recolección de los datos cuantitativos: segunda parte”. Entre tales técnicas se encuentran:

1. Análisis de contenido cuantitativo

Es una técnica para estudiar cualquier tipo de comunicación de una manera “objetiva” y sistemática, que cuantifica los mensajes o contenidos en categorías y subcategorías, y los somete a análisis estadístico.

Sus usos son muy variados, por ejemplo: evaluar el grado de carga de contenido sexual de uno o varios programas televisivos; estudiar las apelaciones y características de campañas publicitarias (digamos, de perfumes femeninos de costo elevado) en los medios de comunicación colectiva (radio, televisión, periódicos y revistas); comparar estrategias propagandísticas de partidos políticos en internet; conocer discrepancias ideológicas entre varios periódicos al tratar un tema como el terrorismo internacional; determinar la evolución de cierta clase de pacientes que asisten a psicoterapia al analizar sus escritos y expresiones verbales; cotejar el vocabulario aprendido por pequeños que se exponen más al uso de la computadora en comparación con niños que la utilizan menos; conocer y contrastar la posición de diversos presidentes latinoamericanos en cuanto al problema del desempleo; comparar estilos de escritores que se señalan como parte de una misma corriente literaria; y/o analizar la calidad y profundidad de la información presente en internet sobre un virus.

Una investigación de este tipo es la de Guillaume y Bath (2008), quienes estudiaron la cobertura y el tratamiento que se daba en la prensa británica a la información sobre las vacunas para el sarampión, las paperas y la rubéola durante un periodo de dos meses. Hall y Wright (2008) aplicaron el análisis de contenido para examinar opiniones judiciales.

2. Observación

Este método de recolección de datos consiste en el registro sistemático, válido y confiable de comportamientos y situaciones observables, a través de un conjunto de categorías y subcategorías. Útil, por

ejemplo, para analizar conflictos familiares, eventos masivos (como la violencia en los estadios de fútbol), la aceptación-rechazo de un producto en un supermercado, el comportamiento de personas con capacidades mentales distintas, etc. Haynes (1978) menciona que es el método más utilizado por quienes se orientan conductualmente.

Como muestras de este tipo de investigación podemos citar a Regina *et al.* (2008), quienes utilizando una técnica conocida como la lista de verificación de la conducta autista, compararon las observaciones de profesionales de la salud en torno a los comportamientos autistas de niños brasileños con las observaciones de sus madres. Asimismo, Franco, Rodrigues y Balcells (2008) evaluaron la pedagogía de instructores de ejercicios físicos y aeróbicos en tres gimnasios de Portugal, al analizar por observación clases grabadas en video. Labus, Keefe y Jensen (2003) revisaron estudios para indagar sobre la relación entre los autorreportes de intensidad del dolor y las observaciones directas de la conducta producida por tal dolor.

3. Pruebas estandarizadas e inventarios

Estas pruebas o inventarios miden variables específicas, como la inteligencia, la personalidad en general, la personalidad autoritaria, el razonamiento matemático, el sentido de vida, la satisfacción laboral, el tipo de cultura organizacional, el estrés preoperatorio, la depresión posparto, la adaptación al colegio, intereses vocacionales, la jerarquía de valores, el amor romántico, la calidad de vida, la lealtad a una marca de algún producto, etc. Hay miles de ellas(os).

Asimismo, hay un tipo de pruebas que evalúan proyecciones de los participantes y determinan su estado en una variable, con elementos cuantitativos y cualitativos: las pruebas proyectivas como el *test de Rorschach* (que presenta manchas de tinta en tarjetas o láminas blancas numeradas a los intervinientes y éstos relatan sus asociaciones e interpretaciones en relación con manchas).

4. Datos secundarios (recolectados por otros investigadores)

Implica la revisión de documentos, registros públicos y archivos físicos o electrónicos. Por ejemplo, si nuestra hipótesis fuera: “la violencia manifiesta en la Ciudad de México es mayor que en la ciudad de Caracas”; entonces acudiríamos a las alcaldías de las ciudades para solicitar datos relacionados con la violencia, como número de asaltos, violaciones, robos a casa-habitación, asesinatos, etc. (datos generales, por distrito y habitante). También obtendríamos información de los archivos de los hospitales y las diferentes procuradurías o cuerpos policíacos. Un caso de una investigación cuyo método de recolección se fundamentó en datos secundarios fue el que a continuación se comenta.

EJEMPLO

Un grupo de investigadores efectuó —en 2008 y principios de 2009— un estudio para explorar el impacto que tienen las becas otorgadas y/o gestionadas por una institución de educación superior sobre el desarrollo académico de los alumnos beneficiarios y su deserción escolar.²⁴

Los investigadores solicitaron a las diferentes direcciones información de los estudiantes respecto a su promedio general en la carrera, nivel socioeconómico, estatus respecto a la beca (becado-no becado), tipo de beca (institucional, otorgada por el Ministerio de Educación, por organismo privado, con fondos del gobierno estatal), monto de la beca, estatus académico del alumno (regular, irregular, desertor), semestre que cursa, género y edad, entre otras cuestiones. Consideraron los últimos cinco años escolares. Con tal información construyeron una base de datos (con más de medio millón de registros) y efectuaron análisis. Entre otras cuestiones encontraron que el promedio de los becarios era muy superior al de los no becarios y la deserción escolar era mínima entre los primeros, casi inexistente. Pero no encontraron una relación entre el monto de la beca y el promedio general de la carrera (acumulado). Asimismo, descubrieron que las mujeres tenían en general mejor promedio que sus compañeros.

²⁴ No se menciona el nombre de la institución porque ésta solicitó el anonimato, tampoco el de todos los investigadores, entre los que se encontraban dos de los autores de la presente obra.

Comparar indicadores económicos de países de la Comunidad Europea, analizar la relación comercial entre dos naciones, cotejar el número y tipo de casos atendidos por diferentes hospitales, contrastar la efectividad con que se insertan en el mundo laboral los egresados de una carrera de distintas universidades, evaluar las tendencias electorales en un país, antes y después de un suceso crítico (como lo fueron los deplorables actos terroristas en Madrid en 2004), son ejemplos donde la recolección y análisis de datos secundarios son la base de la investigación.

5. Instrumentos mecánicos o electrónicos

Sistemas de medición por aparatos, como el detector de mentiras, o polígrafo, que considera la respuesta galvánica de la piel (en investigaciones sobre crímenes); la pistola láser, que mide la velocidad a la que circula un automóvil desde un punto externo al vehículo (en estudios sobre el comportamiento de conductores); instrumentos que captan la actividad cerebral (evaluaciones médicas y psicológicas); el escáner, que mide con exactitud el cuerpo de un ser humano y ubica la talla ideal para confeccionar toda su ropa o vestuario (en investigaciones para diseñar los uniformes de los soldados); la medición eléctrica de distancias, etcétera.

6. Instrumentos específicos propios de cada disciplina

En todas las áreas de estudio se han generado valiosos métodos para recolectar datos sobre variables específicas. Por ejemplo, en la comunicación organizacional se utilizan formatos para evaluar el uso que hacen los ejecutivos de los medios de comunicación interna (teléfono, reuniones, etc.), así como herramientas para conocer procesos de comunicación en la empresa (la auditoría en comunicación). Para el análisis de grupos se usan los sistemas sociométricos y el análisis de redes, entre otros.

¿Puede utilizarse más de un tipo de instrumento de recolección de datos?

Cada día es más común ver estudios donde se utilizan diferentes métodos de recolección de datos. En los estudios cuantitativos no resulta extraño que se incluyan varios tipos de cuestionarios al mismo tiempo que pruebas estandarizadas y recopilación de contenidos para análisis estadístico u observación. Incluso, al utilizar diversos instrumentos se ayuda a establecer la validez de criterio. No solamente se puede, sino que es conveniente, hasta donde lo permita el presupuesto para investigar.

¿Cómo se codifican las respuestas de un instrumento de medición?



Una vez recolectados los datos, éstos deben codificarse. Ya hemos dicho que las categorías de un ítem o pregunta requieren codificarse con símbolos o números; y esto debe hacerse, porque de lo contrario no se efectuaría ningún análisis o sólo se contaría el número de respuestas en cada categoría (por ejemplo, 25 contestaron “sí” y 24 respondieron “no”).²⁵ Comúnmente, el investigador se interesa en realizar análisis más allá de un conteo de casos por categoría, y actualmente los análisis se llevan a cabo por medio de la computadora u ordenador. 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 ir o no precodificadas (incluir la codificación en el instrumento de medición) y que las preguntas abiertas no suelen estar precodificadas.

Los valores perdidos y su codificación

Cuando las personas no responden a un ítem, contestan incorrectamente (por ejemplo, marcan dos opciones, cuando las alternativas eran mutuamente excluyentes) o no puede registrarse la información, se crean una o varias categorías de valores perdidos y se les asignan sus respectivos códigos.



²⁵ En ediciones anteriores este apartado incluyó la codificación relacionada con el análisis de contenido y la observación, ahora tal codificación se comenta al final de estos dos temas en el capítulo 7 del CD anexo: “Recolección de los datos cuantitativos: segunda parte”.

EJEMPLOS

Sí = 1

No = 0

No contestó = 3

Contestó incorrectamente = 4

Sí = 1

No = 0

Valor perdido por diversas razones = 9

Asimismo, tenemos el caso de preguntas que **no** aplican a ciertos participantes, en estas situaciones debe considerarse y codificarse la categoría: “no aplica”. Por ejemplo, si un cuestionario administrado mediante entrevista a mujeres contuviera las siguientes dos preguntas: ¿durante el último mes realizó alguna compra en la tienda de ropa femenina “Ensueños”?²⁶ y ¿me podría indicar qué artículos o prendas compró?, y una respondiente nos contestara a la primera pregunta que “no” (no había comprado en la tienda), anotaríamos esta categoría, y obviamente no haríamos la segunda pregunta, sino que marcaríamos la opción “no aplica” (la pregunta).

Los valores perdidos pueden reducirse con instrumentos que motiven al participante y no sean muy largos, con instrucciones claras y capacitación a los entrevistadores. Un alto número de valores perdidos (más de 10% indica que el instrumento tiene problemas). Lo adecuado es que no supere 5% respecto del total de posibles datos o valores.

En la forma tradicional, la codificación de las respuestas a preguntas o afirmaciones implica cuatro pasos que comentaremos brevemente sólo para que se refuercen algunos conceptos:

1. Establecer los códigos de las categorías o alternativas de respuesta de los ítems o preguntas

Cuando todas las categorías fueron precodificadas y no se tienen preguntas abiertas, este primer paso no es necesario, ya se efectuó. Si las categorías no fueron precodificadas y/o se tienen preguntas abiertas, deben asignarse los códigos o la codificación a todas las categorías de los ítems. Por ejemplo:

Pregunta no precodificada:

¿Practica usted algún deporte por lo menos una vez a la semana?

Sí No

Se codifica:

1 = Sí

0 = No

Frase no precodificada:

“Creo que estoy recibiendo un salario justo por mi trabajo”.

() Totalmente de acuerdo () De acuerdo () Ni de acuerdo ni en desacuerdo
() En desacuerdo () Totalmente en desacuerdo

²⁶ Nombre ficticio.

Se codifica:

- 5 = Totalmente de acuerdo
- 4 = De acuerdo
- 3 = Ni de acuerdo ni en desacuerdo
- 2 = En desacuerdo
- 1 = Totalmente en desacuerdo

El tema sobre la codificación de preguntas abiertas ya se expuso antes.

2. Elaborar el libro de códigos incluyendo todos los ítems, uno por uno

Una vez que están codificadas todas las categorías de los ítems, se procede a elaborar el “libro de códigos”, el cual describe la localización de las variables y los códigos asignados a las categorías *en una matriz o base de datos*. Los elementos comunes de un libro de códigos son: variables de la investigación, preguntas o ítems, categorías, códigos (números o símbolos utilizados para asignarse a las categorías) y número de columna en la matriz de datos a que corresponde cada ítem.

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

1. “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
2. “Los servicios que presta la Dirección General de Impuestos Nacionales son habitualmente muy buenos”.
 - (5) Muy de acuerdo
 - (4) De acuerdo
 - (3) Ni de acuerdo ni en desacuerdo
 - (2) En desacuerdo
 - (1) Muy en desacuerdo
3. “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 tabla 9.10.

Es decir, el libro de códigos es una especie de manual para interpretar la matriz de datos (la cual como veremos más adelante es una matriz en Excel, SPSS —Paquete Estadístico para las Ciencias Sociales—, Minitab o cualquier otro programa similar).

▲ **Tabla 9.10** Ejemplo de un libro o documento de códigos con una escala de actitud tipo Likert (tres ítems)

Variable	Ítem	Categorías	Códigos	Columnas
Actitud hacia la Dirección General de Impuestos Nacionales	Frase 1 (informa)	— Muy de acuerdo	5	1
		— De acuerdo	4	
		— Ni de acuerdo ni en desacuerdo	3	
		— En desacuerdo	2	
		— Muy en desacuerdo	1	
	Frase 2 (servicios)	— Muy de acuerdo	5	2
		— De acuerdo	4	
		— Ni de acuerdo ni en desacuerdo	3	
		— En desacuerdo	2	
		— Muy en desacuerdo	1	
	Frase 3 (deshonestidad)	— Muy de acuerdo	1	3
		— De acuerdo	2	
		— Ni de acuerdo ni en desacuerdo	3	
		— En desacuerdo	4	
		— Muy en desacuerdo	5	

3. Efectuar físicamente la codificación

El tercer paso del proceso es la codificación física de los datos, es decir, llenar la matriz de datos con los valores implicados en las respuestas al instrumento de medición (transferir éstas a la matriz).

Veamos un ejemplo simplificado con la escala Likert de tres ítems aplicada a cuatro individuos (figura 9. 17).

Persona 1

A continuación...	
1. “La Dirección General de Impuestos Nacionales informa oportunamente sobre cómo, dónde y cuándo pagar los impuestos”.	(5) Muy de acuerdo (X) De acuerdo (3) Ni de acuerdo ni en desacuerdo (2) En desacuerdo (1) Muy en desacuerdo
2. “Los servicios que presta la Dirección General de Impuestos Nacionales son en general muy buenos”.	(X) Muy de acuerdo (4) De acuerdo (3) Ni de acuerdo ni en desacuerdo (2) En desacuerdo (1) Muy en desacuerdo
3. “La Dirección General de Impuestos Nacionales se caracteriza por la deshonestidad de sus funcionarios”.	(5) Muy de acuerdo (4) De acuerdo (X) Ni de acuerdo ni en desacuerdo (2) En desacuerdo (1) Muy en desacuerdo
El participante obtuvo:	4 (de acuerdo) 5 (muy de acuerdo) 3 (ni de acuerdo ni en desacuerdo)
<i>Persona 2</i>	
Obtuvo respectivamente:	3 (ni de acuerdo ni en desacuerdo) 4 (de acuerdo) 3 (ni de acuerdo ni en desacuerdo)
<i>Persona 3</i>	
Obtuvo respectivamente:	4 4 4
<i>Persona 4</i>	
Obtuvo respectivamente:	5 4 3

Figura 9.17 Ejemplo de aplicación de tres ítems a cuatro sujetos.

De acuerdo con el libro de códigos de la tabla 9.10 y las respuestas a la escala, tendríamos la matriz de la figura 9.18.

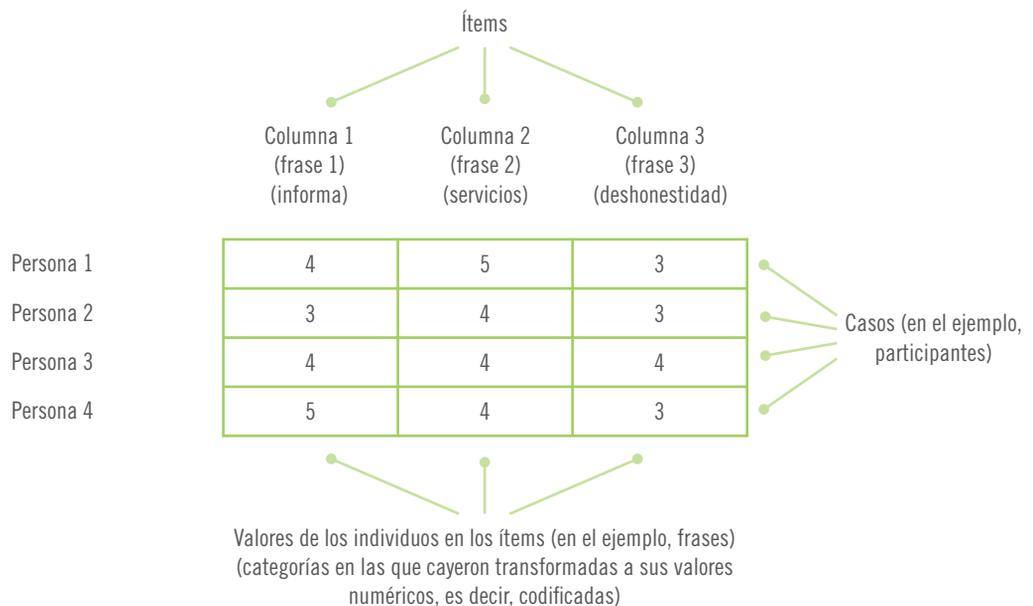


Figura 9.18 Ejemplo de matriz de datos para el libro de códigos de la tabla 9.10.

4. Guardar los datos codificados (casos) en un archivo permanente

Dicho de otro modo, guardar la matriz como documento de SPSS®, Excel, Minitab o equivalente, y por supuesto, darle un nombre que lo identifique.

Codificación utilizando un programa de análisis estadístico



Pero hoy en día los investigadores ya no lo hacen de la manera descrita, sino que la codificación la efectúan directamente, transfiriendo los valores registrados en los instrumentos aplicados (cuestionarios, escalas de actitudes o equivalentes) a un archivo/matriz de un programa computarizado de análisis estadístico (SPSS®, Minitab o equivalente). O bien, si no se cuenta con el programa, los datos se capturan en un documento de Excel (matriz) y luego se trasladan a un archivo del programa de análisis. Veamos el proceso en SPSS®, pero antes es necesario hacer algunas aclaraciones:



- Se abre el programa SPSS®, como cualquier otro, y si se trata de un archivo existente con los datos codificados (matriz completada), pues lo abrimos y hacemos los análisis pertinentes. Si vamos a crear un nuevo archivo o base de datos, elegimos la opción: “teclear datos” y comenzamos a ejercer tal función.
- SPSS® y programas equivalentes tienen dos matrices o ventanas: *a) vista de las variables* (variable view) y *b) vista de los datos* (data view). Ambas aparecen como pestañas (simulando carpetas o folders) ubicadas en la parte inferior de la pantalla hacia nuestro lado izquierdo.
- La “vista de variables” representa el sistema de codificación o libro/documento de códigos electrónico (constituye una matriz). Los renglones o filas significan ítems o reactivos, y las columnas representan características, propiedades o atributos de cada ítem. A los ítems en estos programas se les denomina “variables” de la matriz, a veces coinciden con el concepto de variable que se tiene en la investigación (por ejemplo, género, es una variable de la investigación y un renglón o fila en la “vista de variables”) y en otras ocasiones son simplemente un ítem de una variable del estudio.

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columna	Alineación	Medida
Ítem 1 o variable 1 de la matriz										
Ítem 2 o variable 2 de la matriz										
Ítem k...										

Como dijimos, las columnas son propiedades del ítem que debemos definir:

1. Nombre de cada ítem o variable de la matriz: lo asignamos nosotros (obviamente debe reflejar al ítem o reactivo al que hace referencia). Por ejemplo: género, edad, p1 (pregunta uno), ingresos, etcétera.
2. Tipo de variable de la matriz (numérica, no numérica o cadena —símbolos o números que indican un nivel nominal, como una fecha—, etc.). Incluso la clase puede ser numérica, como una cifra con decimales. Este tipo se vincula al nivel de medición. Asimismo, es necesario especificar el ancho (caracteres) de la variable y los decimales, si tiene (por ejemplo, si la variable implica cantidades en moneda y centavos).
3. Anchura (en dígitos o caracteres). Esto depende de la comodidad de ancho con la cual deseamos trabajar y del ancho de las categorías (ejemplos: en un ítem actitudinal la calificación ocupa un dígito —totalmente de acuerdo = 5, de acuerdo = 4, etc.—, ingresos puede ocupar varios dígitos de acuerdo con el tipo de moneda —si no agrupamos y decidimos colocar la cantidad completa—. El ancho debe coincidir con el especificado en tipo de variable.
4. Decimales (si es pertinente). Es necesario que coincidan con los expresados en tipo de variable.
5. Etiqueta (definición o párrafo que describe a la variable o ítem). Por ejemplo: antigüedad en la empresa, ingresos acumulados en el año, pregunta uno de la prueba sobre inteligencia emocional...
6. Valores. Los códigos de cada opción de respuesta o categoría. La codificación en sí. Incluye, desde luego, valor (por ejemplo = 1) y su etiqueta (“mujer”). También de los valores perdidos.
7. Valores perdidos (se especifican los códigos de las categorías u opciones de los valores perdidos).
8. Columnas (una vez más el número de dígitos que ocupa la variable, contando decimales y el punto decimal, si es pertinente). Debe coincidir con anchura.
9. Alineación (si queremos que los datos, cifras o valores en la matriz o vista de los datos se alineen a la derecha, izquierda o al centro).
10. Medida (nivel de medición del ítem: escala —intervalo o razón—, ordinal o nominal).

En la figura 9.19 se muestra un ejemplo de la vista de las variables en SPSS®.

- La “vista de los datos” es la matriz de datos. Las columnas son ítems o variables de la matriz y los renglones o filas representan casos (unidades, participantes, etc.); mientras que las celdas son los datos o valores. Cada celda representa un valor de un caso en una variable o ítem.

En la figura 9.20 mostramos un ejemplo de la vista de los datos.

Errores de codificación

Al teclear los valores en la vista de los datos, se pueden cometer errores, es humano. Por ejemplo, que en un ítem o variable de la matriz donde solamente se tenían dos categorías, aparezca en uno o más casos una no contemplada (imaginemos que tenemos el ítem *género* con las opciones: 1 = masculino y 2 = femenino, y alguien teclea un “3” o un “8”, esto es un error de codificación; o bien que en una

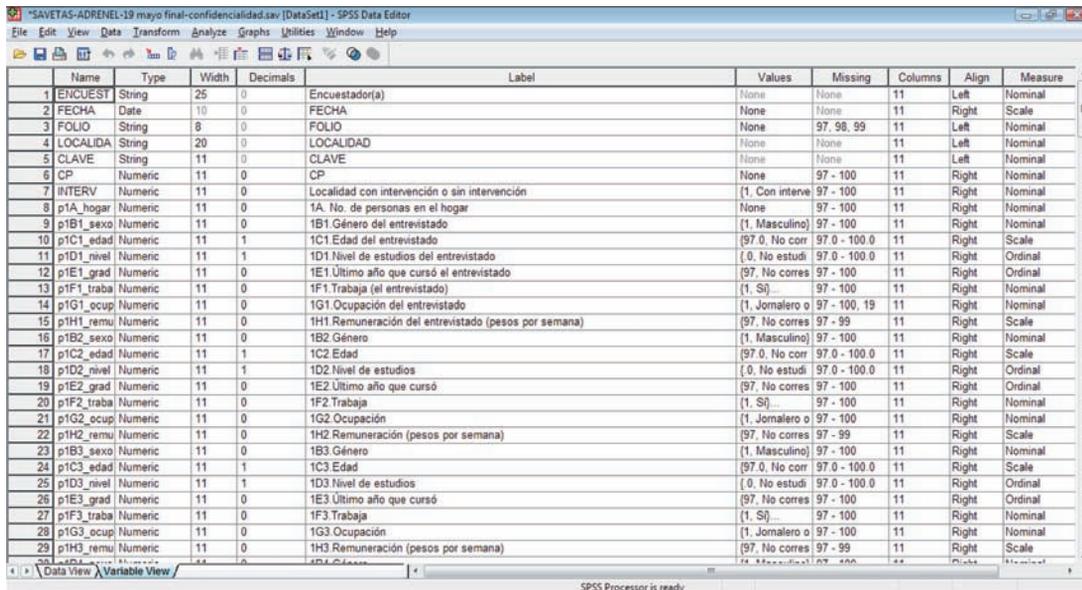


Figura 9.19 Ejemplo de la vista de las variables en SPSS®.

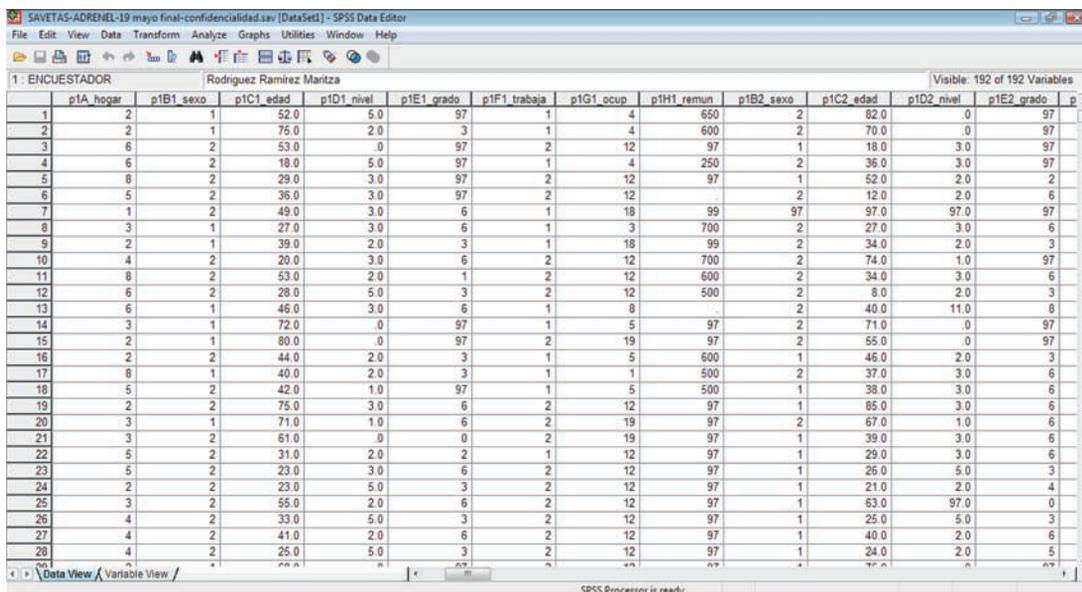


Figura 9.20 Ejemplo de la vista de los datos en SPSS®.

escala con tres opciones: 1. en desacuerdo, 2. neutral y 3. de acuerdo, se presenten valores como: “w”, “#”, ¿qué es eso?). Los errores de codificación tienen que corregirse. Esto puede hacerse: a) revisando físicamente la vista de los datos y haciendo los cambios pertinentes, b) en SPSS® con la función “ordenar o clasificar casos” —sort— (en “Datos” o “Data”) y, de este modo, visualizar valores que no correspondan a cada variable o ítem de la matriz, c) ejecutando el “análisis de frecuencias” en el menú “Analyze” y “Estadísticos descriptivos”, y una vez obtenidos los resultados, se observará en qué variables de la matriz (columnas) hay valores que no deberían estar, para efectuar las correcciones necesarias.²⁷

²⁷ Desde luego, como SPSS® se actualiza permanentemente los comandos pueden variar, mas no las funciones.

Cabe señalar que los *valores perdidos* no son errores de codificación, porque al registrarlos como tales, le estamos informando al programa que son precisamente valores perdidos y si lo deseamos, se excluyen del análisis de frecuencias, salvo que queramos saber cuántos no contestaron o lo hicieron incorrectamente (pero por default no cuentan, por ejemplo, para calcular promedios y análisis inferenciales).

En ambas vistas se muestran las opciones para ejecutar las funciones de SPSS® (más recientemente denominado PASW Statistics Base), como por ejemplo: analizar datos y elaborar gráficas, las cuales se comentarán en el siguiente capítulo, el 10: “Análisis de los datos cuantitativos” y con mayor profundidad en el CD anexo: Material complementario → Manuales → Manual “Introducción al SPSS®/PASW”. Este manual lo llevará por el proceso. Además, hay cientos de páginas sobre este paquete y la de la propia empresa a nivel mundial (<https://www.spss.com/>).²⁸

En el mismo CD se encuentra un demo de este programa. Entonces, el proceso sería el que se muestra en la figura 9.21.

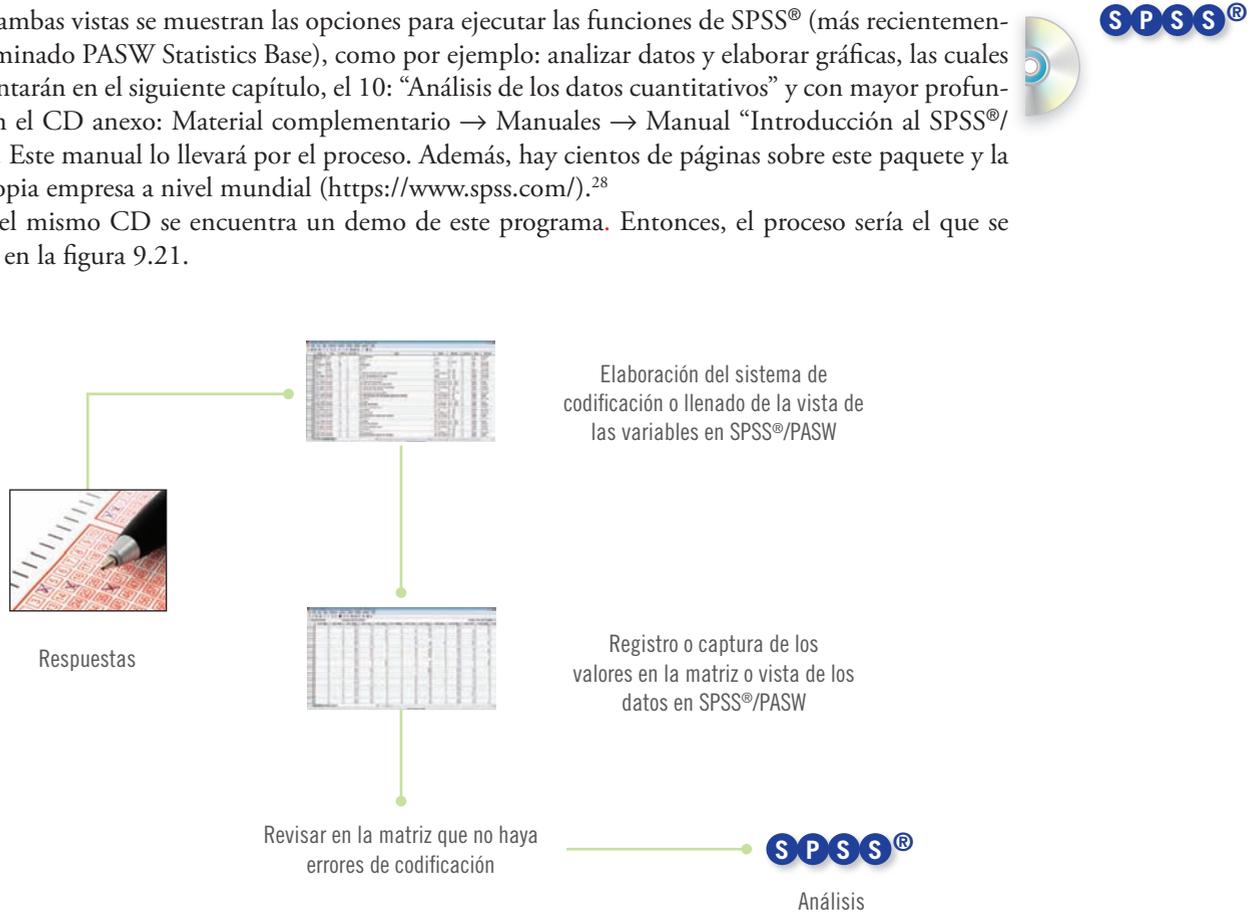


Figura 9.21 Proceso de codificación y preparación de los datos para su análisis en SPSS®.

Si queremos capturar los datos en nuestra PC u ordenador personal y no disponemos de SPSS®/PASW (solamente en nuestra universidad, centro de cómputo público o empresa), podemos hacerlo en una matriz de Excel y luego copiarlos y pegarlos en la vista de los datos de SPSS®/PASW. Pues las columnas (A, B, C, D, etc.) corresponden a las variables de la matriz, y los renglones o filas, son los casos, al igual que en SPSS®/PASW.

Como todo archivo, debe guardarse y respaldarse, implica nuestros datos y el sistema de codificación.

Cuando se utilizan dispositivos electrónicos para capturar los datos (como palms, lectores ópticos, cuestionarios electrónicos), obviamente no requerimos teclear los datos, éstos pasan directamente de la fuente a la matriz o base de datos.

En este capítulo, por razones didácticas, se presentaron matrices pequeñas, pero en la investigación pueden tenerse de 500 o más columnas.

²⁸ Asimismo, busque en su país o región al representante de SPSS® Inc.



Resumen

- Recolectar los datos implica: *a)* seleccionar uno o varios métodos o instrumentos disponibles, adaptarlo(s) o desarrollarlo(s), esto depende del enfoque que tenga el estudio, así como del planteamiento del problema y de los alcances de la investigación; *b)* aplicar el (los) instrumento(s), y *c)* preparar las mediciones obtenidas o los datos recolectados para analizarlos correctamente.
- En el enfoque cuantitativo, recolectar los datos es equivalente a medir.
- Medir es el proceso de vincular conceptos abstractos con indicadores empíricos, mediante clasificación o cuantificación.
- En toda investigación cuantitativa medimos las variables contenidas en la(s) hipótesis.
- Cualquier instrumento de recolección de datos debe cubrir dos requisitos: confiabilidad y validez.
- La confiabilidad se refiere al grado en que la aplicación repetida de un instrumento de medición, a los mismos individuos u objetos, produce resultados iguales.
- La validez se refiere al grado en que un instrumento de medición mide realmente la(s) variable(s) que pretende medir.
- Se pueden aportar tres tipos principales de evidencia para la validez cuantitativa: 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: la improvisación, utilizar instrumentos desarrollados en el extranjero y que no han sido validados para nuestro contexto, poca o nula empatía con los participantes y los factores de aplicación.
- No hay medición perfecta, pero el error de medición debe reducirse a límites tolerables.
- La confiabilidad cuantitativa se determina al calcular el coeficiente de fiabilidad.
- Los coeficientes de fiabilidad cuantitativa varían entre 0 y 1 (0 = nula confiabilidad, 1 = total confiabilidad).
- Los métodos más conocidos para calcular la confiabilidad son: *a)* medida de estabilidad, *b)* formas alternas, *c)* mitades partidas, *d)* consistencia interna.
- La evidencia sobre la validez de contenido se obtiene al contrastar el universo de ítems frente a los ítems presentes en el instrumento de medición.
- La evidencia sobre la validez de criterio se obtiene al comparar los resultados de aplicación del instrumento de medición frente a los resultados de un criterio externo.
- La evidencia sobre la validez de constructo se puede determinar mediante el análisis de factores y al verificar la teoría subyacente.
- Los pasos genéricos para elaborar un instrumento de medición son:
 1. Redefiniciones fundamentales sobre propósitos, definiciones operacionales y participantes.
 2. Revisar la literatura, particularmente la enfocada en los instrumentos utilizados para medir las variables de interés.
 3. Identificar el conjunto o dominio de conceptos o variables a medir e indicadores de cada variable.
 4. Tomar decisiones en cuanto a: tipo y formato; utilizar uno existente, adaptarlo o construir uno nuevo, así como el contexto de administración.
 5. Construir el instrumento.
 6. Aplicar la prueba piloto (para calcular la confiabilidad y validez iniciales).
 7. Desarrollar su versión definitiva.
 8. Entrenar al personal que va a administrarlo.
 9. Obtener autorizaciones para aplicarlo.
 10. Administrar el instrumento.
 11. Preparar los datos para el análisis.
- En la investigación social disponemos de diversos instrumentos de medición.
 1. Principales escalas de actitudes: Likert, diferencial semántico y escalograma de Guttman (este último se encuentra comentado en el capítulo 7 del CD anexo).
 2. Cuestionarios (autoadministrado, por entrevista personal, por entrevista telefónica, internet y por correo).
 3. Recolección de contenidos para análisis cuantitativo (capítulo 7 del CD anexo).
 4. Observación cuantitativa (capítulo 7 del CD anexo).
 5. Pruebas estandarizadas (capítulo 7 del CD anexo).
 6. Archivos y otras formas de medición (capítulo 7 del CD anexo).
- Las respuestas a un instrumento de medición se codifican.
- Actualmente, la codificación se efectúa transfiriendo los valores registrados en los instrumentos aplicados (cuestionarios, escalas de actitudes o equivalentes) a un archivo/matriz de un programa computarizado de análisis estadístico (SPSS®, Minitab o equivalente).
- Para resumir algunos de los instrumentos tratados en el capítulo se agraga la tabla 9.11:

▲ **Tabla 9.11** Concentrado de instrumentos para la recolección de datos

Métodos	Propósito general básico	Ventajas	Retos
Cuestionarios/ Escalas de actitudes/ Pruebas estandarizadas	<ul style="list-style-type: none"> • Obtener de manera relativamente rápida datos sobre las variables. • Propios para actitudes, expectativas, opiniones y variables que pueden medirse mediante expresiones escritas o que el mismo participante puede ubicarse en las categorías de las variables (autoubicación). 	<ul style="list-style-type: none"> • Puede ser anónimo. • Poco costosa su aplicación individual. • Relativamente fácil de responder. • Relativamente fácil de analizar y comparar. • Puede administrarse a un considerable número de personas. • Normalmente disponemos de versiones previas para escoger o basarnos en éstas. 	<ul style="list-style-type: none"> • Regularmente no se obtiene retroalimentación detallada de parte de los intervinientes. • Se evalúan actitudes y proyecciones, no comportamientos (mediciones indirectas). • El manejo del lenguaje puede ser una fuente de sesgos e influir en las respuestas. • Son impersonales. • No nos proporcionan información sobre el individuo, excepto en las variables medidas.
Observación	<ul style="list-style-type: none"> • Recolectar información no obstrusiva respecto a conductas y procesos. 	<ul style="list-style-type: none"> • Se puede adaptar a los eventos tal y como ocurren. • Se evalúan hechos, comportamientos y no mediciones indirectas. 	<ul style="list-style-type: none"> • Dificultad para interpretar conductas. • Complejidad al categorizar las conductas observadas. • Puede ser obstrusiva y provocar sesgos si es "participante". • Puede ser costosa.
Análisis de contenido	<ul style="list-style-type: none"> • Recolectar información no obstrusiva respecto de mensajes. 	<ul style="list-style-type: none"> • Se puede adaptar a los eventos tal como ocurren. • Se evalúan mediciones indirectas. 	<ul style="list-style-type: none"> • Dificultad para interpretar mensajes. • Complejidad al categorizar los mensajes.



Conceptos básicos



Análisis cuantitativo de contenido (capítulo 7 del CD anexo)

Archivo de datos

Categorías

Codificación

Codificador

Coeficiente alfa de Cronbach

Coeficiente de confiabilidad

Coeficiente KR-20 de Kuder-Richardson

Confiabilidad

Contexto de administración del instrumento

Cuestionarios

Diferencial semántico

Escala Likert

Escalas de actitudes

Escalograma de Guttman (capítulo 7 del CD anexo)

Entrevista

Evidencia relacionada con el constructo

Evidencia relacionada con el contenido

Evidencia relacionada con el criterio

Hojas de codificación

Instrumento de medición

Matriz de datos

Medición

Medida de estabilidad

Método de formas alternas

Método de mitades partidas

Niveles de medición

Observación cuantitativa (capítulo 7 del CD anexo)

Pruebas estandarizadas (capítulo 7 del CD anexo)

Pruebas proyectivas (capítulo 7 del CD anexo)

Recolección de datos

Registro de codificación (capítulo 7 del CD anexo)

Unidad de análisis

Validez

Vista de las variables

Vista de los datos





Ejercicios



1. Busque una investigación cuantitativa en algún artículo de una revista científica, en la cual 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 explique con ejemplos la diferencia entre confiabilidad y validez.
3. Defina ocho variables e indique su nivel de medición.
4. Defina una variable de cada nivel de medición.
5. Suponga que alguien intenta evaluar la actitud hacia el Presidente de la República, entonces construya un cuestionario tipo Likert con 20 ítems para medir dicha actitud, e indique cómo se calificaría la escala total (10 ítems positivos y 10 negativos). Por último, señale la dimensión que cada ítem pretende medir de dicha actitud (credibilidad, presencia física, etcétera).
6. Construya un cuestionario para medir la variable que considere conveniente (con un mínimo de 10 preguntas o ítems) e incluya preguntas demográficas adicionales. Aplíquelo a 20 conocidos suyos; elabore el libro de códigos y la matriz de datos, mínimo en Excel.
7. ¿Cómo mediría la hostilidad mediante observación y cómo por medio de una escala de actitudes? (Debe leer el apartado de observación en el capítulo 7 del CD anexo.)
8. Genere un planteamiento del problema, donde utilice por lo menos dos tipos de instrumentos cuantitativos para recolectar datos.
9. ¿Cómo se podría aplicar el análisis cuantitativo del contenido para la evaluación de un programa educativo a nivel superior? (Debe leer el apartado de análisis de contenido en el capítulo 7 del CD anexo).
10. Construya una matriz de datos sobre las siguientes variables: género, edad, deporte preferido para practicar, deporte preferido para observar, escuela de procedencia (pública/privada), tipo de música que más le agrada, si está o no en desacuerdo con la política económica del gobierno actual, partido por el que votó en las últimas elecciones municipales y líder histórico que más admira. Que los participantes sean sus compañeros de clase (el ejercicio implica levantar datos y codificarlos, desde luego).
11. Elabore uno o varios instrumentos para el ejemplo de estudio que ha desarrollado hasta ahora en el proceso cuantitativo (incluya la codificación).



Ejemplos desarrollados²⁹

La televisión y el niño

Se aplicó un cuestionario en una muestra total de 2112 niños y niñas del Distrito Federal (capital de México), de acuerdo con la estrategia de muestreo planteado. Las variables medidas fueron: uso de medios de comunicación colectiva, tiempo de exposición a la televisión, preferencia de contenidos televisivos (programas), bloques de horarios de exposición a la televisión (mañana, tarde y/o noche), comparación de la televisión con otras fuentes de entretenimiento, actividades que realiza mientras observa la televisión, condiciones de exposición a la televisión (solo-acompañado), autonomía en la elección de los programas, control de los padres sobre la actividad de ver televisión, usos y gratificaciones de la televisión, datos demográficos.



El cuestionario es descriptivo y fue explorado y validado por 10 expertos en investigación sobre la relación niño-televisión. Se elaboró una versión piloto con 100 infantes (50 niñas y 50 niños), la cual se probó y ajustó. No hubo escalas con varios ítems, por lo que no se calcularon coeficientes de confiabilidad.

La pareja y relación ideales

Se desarrolló un instrumento inicial para recolectar los datos con base en preguntas. Por ejemplo: “pensando en tu relación ideal, ¿cuáles serían las características que más te gustaría que tuviera esa relación?” o bien: “¿qué cualidades te gustaría que tuviera tu novio ideal?” Sin embargo, en la prueba piloto con 100 estudiantes se observó que era mejor sustituir las preguntas por afirmaciones que

²⁹ Por cuestiones de espacio, se comentan brevemente. El primer ejemplo aborda un aspecto de la recolección: el procedimiento y las variables centrales; el segundo, el instrumento de medición; y el tercero, la consideración y comparación de instrumentos (escalas).

fueran completadas por los participantes (por ejemplo: “pensando en tu relación ideal, las características que más te gustaría que tuviera esa relación serían...” Resultaban más comprensibles para ellos y fueron respondidas con mayor precisión. Así, se aplicó por entrevista el siguiente cuestionario:³⁰

Cuestionario sobre la pareja y relación ideales

El objetivo de esta encuesta es conocer tu opinión acerca de las relaciones y parejas sentimentales que has tenido, tienes y tendrás, así como de tu concepción de una pareja ideal, por lo cual te agradeceríamos contestar lo que se te indica a continuación, pensando y contestando según corresponda en cada caso.

Recuerda que tus respuestas son totalmente confidenciales.

Edad: _____ años

Género: 1) Masculino _____ 2) Femenino _____

Indica la carrera que estudias actualmente:

Psicología ()	Turismo ()	Comunicación ()
Medicina ()	Administración ()	Mercadotecnia ()
Arquitectura ()	Contaduría Pública ()	Derecho ()
Ing. Industrial ()	Ing. en Sistemas Computacionales ()	Comercio Internacional ()
Otra (mencionar) _____		

1. Para ti ¿qué es un novio(a)? Un novio(a) es: _____

2. Para ti ¿qué es una relación de noviazgo? Un no-viazgo es: _____

 _____.

Pasado:

3. ¿Has tenido novio(a)? 1) Sí _____ 2) No _____
4. ¿Con cuántos(as) novios(as) has durado más de un mes? _____.
5. Las cualidades que más te gustaban del novio(a) más importante que has tenido en el pasado son:
 Anota de la más importante 1) a la menos importante 5).

1) _____, 2) _____,
 3) _____, 4) _____,
 5) _____.

6. Pensando en tu relación pasada más importante, las características que más te gustaban de la relación de pareja eran (no hablamos de tu pareja, sino de la relación de noviazgo):

Anota de la más importante 1) a la menos importante 5).

1) _____, 2) _____,
 3) _____, 4) _____,
 5) _____.

Actualmente:

7. ¿Tienes novio(a)? 1) Sí _____ 2) No _____

8. Las cualidades que más te gustan de tu novio(a) son:

Anota de la más importante 1) a la menos importante 5).

1) _____, 2) _____,
 3) _____, 4) _____,
 5) _____.

9. ¿Cuántos meses llevas con tu novio(a) actual? _____

10. Pensando en tu relación actual, las características que más te gustan de la relación de pareja (no hablamos de tu pareja, sino de la relación de noviazgo) son:

Anota de la más importante 1) a la menos importante 5).

1) _____, 2) _____,
 3) _____, 4) _____,
 5) _____.

11. ¿Qué tan importante es en tu vida tu pareja actual?

1) Sumamente importante 2) Muy importante
 3) Importante 4) Poco importante 5) No tiene importancia

Ideal:

12. Piensa en tu novio(a) ideal y menciona las cualidades que te gustaría que tuviera:

Anota de la más importante 1) a la menos importante 5).

1) _____, 2) _____,
 3) _____, 4) _____,
 5) _____.

13. Pensando en tu relación ideal, las características que más te gustaría que tuviera esa relación

³⁰ Las opciones de respuesta se han reducido, también por espacio (por ejemplo: carreras).

(no hablamos de tu pareja, sino de la relación de noviazgo) serían:

Anota de la más importante 1) a la menos importante 5).

- 1) _____, 2) _____,
3) _____, 4) _____,
5) _____.

Futuro:

14. En tu futuro, ¿te gustaría o no tener una relación de pareja para toda la vida?
1. Sí ____ 2. No ____ 3. No sé ____
15. ¿Por qué? _____.
16. En tu futuro ¿qué tipo de relación de pareja duradera a largo plazo te gustaría establecer, tener o formar?
1. Matrimonio civil.
2. Matrimonio religioso.
3. Matrimonio religioso y civil.
4. Unión libre (vivir juntos sin estar casados).
5. Llevar una relación de pareja sin vivir juntos.
6. Otra: _____.

Gracias por tu colaboración.

El abuso sexual infantil

Escala cognitiva

El instrumento Children's Knowledge of Abuse Questionnaire-Revised (CKAQ-R), fue traducido al español y adaptado para preescolares. En esta escala adaptada, se eliminaron los elementos redundantes y los que evaluaban las actitudes ante los desconocidos, bajo la tesis que quienes agreden sexualmente a los menores son en su gran mayoría personas cercanas. Además, se simplificaron las preguntas formuladas negativamente, tales como "¿algunas veces está bien no hacer lo que nos pide un adulto?", que tienden a ser confusas para los preescolares. El CKAQ-Español puede tener un puntaje máximo de 22, cada reactivo posee evaluación dicotómica, dando un punto por cada respuesta correcta. Sigue el mismo esquema y protocolo que el CKAQ original. Cada pregunta puede ser contestada como "sí", "no", o "no sé" y su evaluación es dicotómica (correcto o incorrecto). Incluye cuestiones para medir el desarrollo cognitivo y actitudes asertivas ante contactos positivos y negativos, chantaje emocional, disociación de los contactos con la afectividad y el pedir ayuda ante el abuso.

El estudio de la confiabilidad interna se efectuó con el modelo Kurder Richardson 20 (KR-20), bajo la versión adaptada de Cronbach para reactivos dicotómicos. Tal estudio se realizó con el total de

casos ($n = 150$). Se obtuvo un alfa de 0.69, lo que representa un nivel moderadamente aceptable.

Escala conductual

Después del estudio de diversas escalas, se decidió partir del RPP para el desarrollo del instrumento conductual. Las razones para esta decisión se basaron en que el RPP se ha aplicado a muestras grandes ($n = 670$). Por otro lado, evalúa en acción los patrones seguidos por los agresores, así da la oportunidad de analizar las reacciones de los niños y sus habilidades de protección "en vivo". Además, no aborda al niño o niña de manera burda o aterradora, se enfoca en los preámbulos del abuso, en donde se censa la posibilidad. Por estas razones, este instrumento nos parece de los más acertados por su evaluación conductual, su aproximación a lo que un(a) niño(a) puede vivir en su cotidianidad en cuanto a sus aproximaciones incómodas y evaluar sus recursos asertivos, seguridad emocional y habilidades de autoprotección.

Uno de los inconvenientes de este protocolo es que no se disponen de valores psicométricos que lo avalen. Por lo que no hay comparativos para los resultados que de esta investigación se obtengan.

Partiendo del RPP original, se le hizo una adaptación mediante la traducción y adecuación al contexto mexicano. A esta escala le llamaremos Role Play-México. Uno de los inconvenientes que se le cuestiona al RPP es que sólo puede aplicarse uno a uno. Es decir, no se puede aplicar a grupos de infantes en conjunto. Sin embargo, en el caso de preescolares esto no aplica, porque en general las pruebas administradas a grupos, requieren del desarrollo de las habilidades lectoescritoras, un estado no dominado en la etapa preescolar. Por tanto, tal inconveniente es intrascendente en el caso de estudio. Otra desventaja que se le atribuye, es que la escala no incluye elementos que evalúen la actitud de los menores ante los contactos positivos, para determinar si los PPASI (programas de prevención de abuso sexual infantil) generan un efecto nocivo de suspicacia indiscriminada ante cualquier contacto. Por lo cual, se decidió incluir un par de reactivos para evaluar esta posibilidad en el Role Play-México (RPMéxico). Estos reactivos incluyen por ejemplo, abrazos por los padres o felicitaciones. Se desarrolló, también, una prueba paralela a dicha adaptación, a la que llamamos Evaluación de la Prevención del Abuso (EPA).

En la escala RPP se tiene un puntaje máximo de 14 puntos, evalúa la negación verbal y no verbal de seis escenas "en vivo". Es decir, donde el evaluador actúa y se le pide al niño que responda a la pregunta: ¿qué diría y haría? en una situación planteada. Además, en los tres reactivos donde se abor-

da el chantaje emocional y la coerción, se otorga un punto extra si el participante muestra intención de denunciar el evento. En el caso de la evaluación del RPMéxico y de la EPA se considera un total de ocho escenarios “en vivo”, seis de tipo abusivo y dos de contactos no abusivos. El puntaje de ambas escalas (RPMéxico y EPA) tienen un máximo de 40 puntos. Al igual que el RPP, evalúa la asertividad verbal y conductual, pero se amplía la evaluación con la intención de denuncia del evento abusivo, cubriendo la necesidad de mejorar el sistema de medición con la persistencia de los infantes de pedir ayuda hasta obtenerla. Mide además las siguientes subescalas: 1) reconocimiento de contactos, tanto positivos como negativos, y las habilidades de asertividad verbal (*qué decir*), no verbal (*qué hacer*) y la persistencia en la intención de denuncia ante algún incidente abusivo (*denuncia*). Los 40 puntos, se derivan de la suma de un punto por cada acierto en la asertividad verbal (ocho máximo), un punto por cada asertividad conductual (máximo ocho), un punto por cada intención de denuncia de los contac-

tos inapropiados (seis máximo) y un punto por cada persona a quien denunciarían, hasta un máximo de tres por cada escenario de contacto inapropiado (18 puntos máximo).

Se desarrolló el análisis de confiabilidad tanto temporal como interno. Se aplicó *test-retest* de acuerdo con un método de formas paralelas al administrarse el RPMéxico y el EPA. La correlación entre ambas pruebas alcanzó un buen nivel y fue significativo ($r = 0.75$, $p < 0.01$), lo que avala la utilización de estos instrumentos de forma paralela. El *test-retest* se aplicó en un subgrupo ($n = 44$) del grupo control ($n = 79$). Este estudio confirma que hay correlación entre *test* y *retest* entre cada instrumento RPP, RPMéxico y en todas las subescalas, los índices van de 0.59 a 0.78, todas con $p < 0.01$. El instrumento RPMéxico tiene una correlación ($r = 0.75$) equivalente a la reportada en otros instrumentos similares (WIST, PSQ). Este índice muestra un grado de estabilidad temporal aceptable, dado el tamaño de la muestra.



Los investigadores opinan

Recolección de datos cuantitativos

Dentro del modelo de investigación cuantitativa, la etapa de recolección de los datos resulta de vital importancia para el estudio, de ella dependen tanto la validez interna como externa.

La validez interna de una investigación depende de una adecuada selección o construcción del instrumento con el cual se va a recolectar la información deseada, la teoría que enmarca el estudio tiene que conjugar perfectamente con las características teóricas y empíricas del instrumento; si esto no ocurre, se corre el riesgo de recolectar datos que a la postre pueden ser imposibles de ser interpretados o discutidos, la teoría y los datos pueden caminar por distintas direcciones. Un ejemplo muy sencillo para graficar este problema sería hacer hipótesis y teorizar en torno a la personalidad sobre la base de una de las teorías de los rasgos y usar un instrumento proyectivo para recolectar los datos. Lo correcto sería que la misma teoría sustente los planteamientos hipotéticos y teóricos, así como fundamente el instrumento. Si bien el ejemplo puede resultar un tanto simple y grosero, en el nivel de las investigaciones de pregrado, este problema resulta bastante común y le es muy difícil manejarlo al estudiante promedio.

Del mismo modo, la recolección de los datos se relaciona con la validez externa del estudio, por

cuanto la generalización depende de la calidad y cantidad de los datos que recolectamos. Por ello, en estudios cuantitativos resulta importante determinar una muestra adecuada, que tenga representatividad en el tamaño y que a la vez refleje la misma estructura existente en la población. Sin una buena muestra de datos, no se puede generalizar; y si se corre este riesgo, el investigador podría llevar sus conclusiones más allá de la realidad, cuando lo que se desea es reflejar la realidad.

Una idea clave, para no tropezar con asuntos insalvables en este momento de la investigación o para no tomar decisiones que conduzcan al error, es hacer un buen proyecto de investigación. En la etapa de la planificación debe quedar claramente establecido y justificado qué instrumento se va a utilizar; cómo, dónde y a quiénes se les aplicará; qué instrucciones se les va a brindar a los sujetos o participantes; qué datos son los que se someterán a tratamiento y cuáles otros no serán tomados en cuenta; cómo se van a tratar los mismos y cómo se llegará desde los datos a la teoría.

EDWIN SALUSTIO SALAS BLAS

Universidad de Lima

Perú

Proceso de investigación
cuantitativa**Paso 9 Analizar los datos**

- Decidir el programa de análisis de datos que se utilizará.
- Explorar los datos obtenidos en la recolección.
- Analizar descriptivamente los datos por variable.
- Visualizar los datos por variable.
- Evaluar la confiabilidad, validez y objetividad de los instrumentos de medición utilizados.
- Analizar e interpretar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial).
- Realizar análisis adicionales.
- Preparar los resultados para presentarlos.

Oa Objetivos del aprendizaje

Al terminar este capítulo, el alumno será capaz de:

- 1 Revisar el proceso para analizar los datos cuantitativos.
- 2 Reforzar los conocimientos estadísticos fundamentales.
- 3 Comprender las principales pruebas o métodos estadísticos desarrollados, así como sus aplicaciones y la forma de interpretar sus resultados.
- 4 Analizar la interrelación entre distintas pruebas estadísticas.
- 5 Diferenciar la estadística descriptiva y la inferencial, la paramétrica y la no paramétrica.

Síntesis

En el capítulo se presentan brevemente los principales programas computacionales de análisis estadístico que emplea la mayoría de los investigadores, así como el proceso fundamental para efectuar análisis cuantitativo. Asimismo, se comentan, analizan y ejemplifican las pruebas estadísticas más utilizadas. Se muestra la secuencia de análisis más común, incluyendo estadísticas descriptivas, análisis paramétricos, no paramétricos y multivariados. En la mayoría de estos análisis, el enfoque del capítulo se centra en los usos y la interpretación de los métodos, más que en los procedimientos de cálculo, debido a que en la actualidad los análisis se realizan con ayuda de una computadora.

Se realiza mediante programas computacionales como:

- SPSS®
- Minitab
- SAS
- STATS

Análisis de datos cuantitativos

Cuyo procedimiento es:

Fases

1. Seleccionar el programa estadístico para el análisis de datos
2. Ejecutar el programa
3. Explorar los datos: analizarlos y visualizarlos por variable del estudio
4. Se evalúa la confiabilidad y validez del o de los instrumentos escogidos
5. Se lleva a cabo análisis estadístico descriptivo de cada variable del estudio
6. Se realizan análisis estadísticos inferenciales respecto a las hipótesis planteadas
7. Se efectúan análisis adicionales
8. Se preparan los resultados para presentarlos

El análisis se realiza tomando en cuenta los niveles de medición de las variables y mediante la estadística, que puede ser

Descriptiva

Distribución de frecuencias

Medida de tendencia central

Media
Mediana
Moda

Medidas de variabilidad

Rango
Desviación estándar
Varianza

Gráficas

Puntuaciones z (en CD anexo)

Inferencia

- Sirve para estimar parámetros y probar hipótesis
- Se basa en la distribución muestral

Análisis paramétrico

- Coeficientes de correlación
- Regresión lineal
- Prueba t
- Prueba de la diferencia de proporciones
- Análisis de varianza
- Análisis de covarianza (en CD anexo)

Análisis no paramétrico

- χ^2 cuadrada
- Coeficientes de Spearman y Kendall
- Coeficientes para tabulaciones cruzadas

Análisis multivariados (en CD anexo)

Nota: Este capítulo se complementa con uno del CD anexo: Material complementario → Capítulos → Capítulo 8 “Análisis estadístico: segunda parte”. Asimismo, con el documento incluido en el CD: “Fórmulas estadísticas”, que contiene las fórmulas que estaban en la parte impresa de ediciones anteriores y el apéndice 4 “Tablas anexas”.



¿Qué procedimiento se sigue para analizar cuantitativamente los datos?

Una vez que los datos se han codificado, transferido a una matriz, guardado en un archivo y “limpiado” de errores, el investigador procede a analizarlos.

En la actualidad, el análisis cuantitativo de los datos se lleva a cabo *por computadora u ordenador*. Ya casi nadie lo hace de forma manual ni aplicando fórmulas, en especial si hay un volumen considerable de datos. Por otra parte, en la mayoría de las instituciones de educación media y superior, centros de investigación, empresas y sindicatos se dispone de sistemas de cómputo para archivar y analizar datos. De esta suposición parte el presente capítulo. Por ello, se centra en la *interpretación de los resultados de los métodos de análisis cuantitativo* y no en los procedimientos de cálculo.

QA1

El análisis de los datos se efectúa sobre la *matriz de datos* utilizando un *programa computacional*. El proceso de análisis se esquematiza en la figura 10.1. Posteriormente veremos paso a paso el proceso.

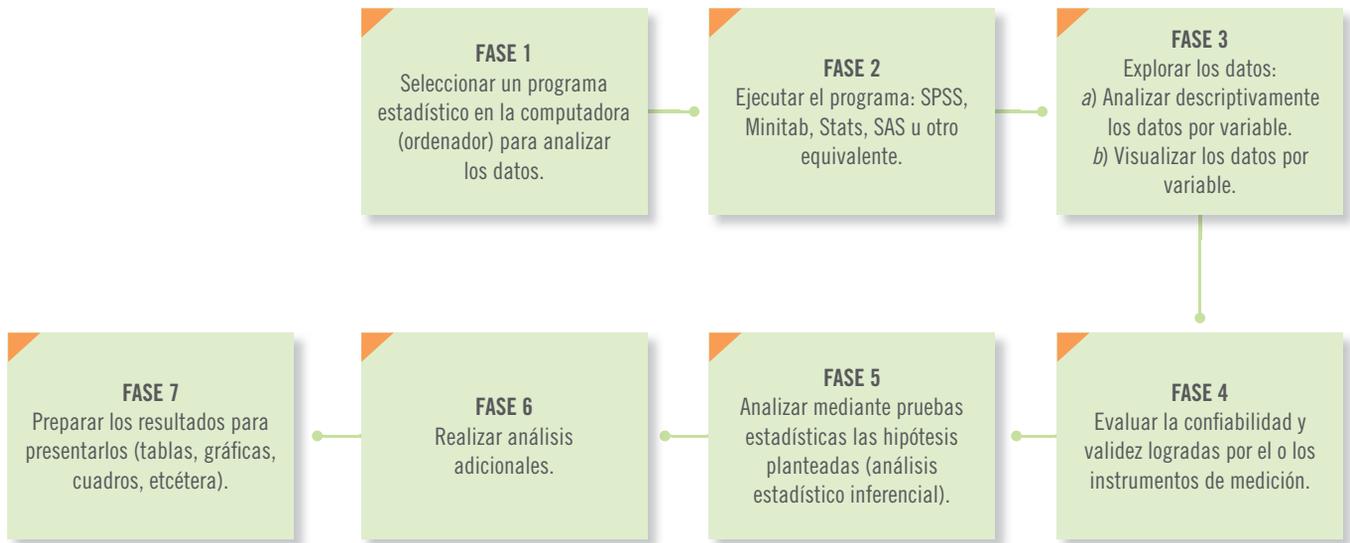


Figura 10.1 Proceso para efectuar análisis estadístico.

Paso 1: seleccionar un programa de análisis

Existen diversos programas para analizar datos. En esencia su funcionamiento es muy similar, incluyen dos partes o segmentos que se mencionaron en el capítulo anterior: una parte de definiciones de las variables, que a su vez explican los datos (los elementos de la codificación ítem por ítem), indicador por indicador en casos propios de las ingenierías y diversas disciplinas y la otra parte, la matriz de datos. La primera parte es para que se comprenda la segunda. Las definiciones, desde luego, son efectuadas por el investigador. Lo que éste hace, una vez recolectados los datos, es precisar los parámetros de la matriz de datos en el programa (nombre de cada variable en la matriz —que equivale a un ítem, reactivo, categoría o subcategoría de contenido u observación, indicador—, tipo de variable o ítem, ancho en dígitos, etc.) e introducir los datos en la matriz, la cual es como cualquier hoja de cálculo. Asimismo, recordemos que la matriz tiene columnas (variables o ítems), filas o renglones (casos) y celdas (intersección entre una columna y un renglón). Cada celda contiene un dato (que significa un valor de un caso en una variable). Supongamos que tenemos cuatro casos o personas y tres variables (género, color de cabello y edad); la matriz se vería como se muestra en la tabla 10.1.

▲ **Tabla 10.1** Ejemplo de matriz de datos con tres variables y cuatro casos

Caso	Columna 1 (género)	Columna 2 (color de pelo)	Columna 3 (edad)
1	1	1	35
2	1	1	29
3	2	1	28
4	2	4	33

La codificación (especificada en la parte de las definiciones de las variables o columnas que corresponden a ítems) sería:

- Género (1 = masculino y 2 = femenino).
- Color de cabello (1 = negro, 2 = castaño, 3 = pelirrojo, 4 = rubio).
- Edad (dato “bruto o crudo” en años).

De esta forma, si se lee por renglón o fila (caso), de izquierda a derecha, la primera celda indica un hombre (1); la segunda, de cabello negro (1), y la tercera, de 35 años (35). En el segundo, un hombre de cabello negro y 29 años. La tercera, una mujer de cabello color negro, con 28 años. La cuarta fila (caso número cuatro) nos señala una mujer (2), rubia (4) y de 33 años (33). Pero, si leemos por columna o variable de arriba hacia abajo, tendríamos en la primera (género) dos hombres y dos mujeres (1, 1, 2, 2).

Por lo general, en la parte superior de la matriz de datos aparecen las opciones de los comandos para operar el programa de análisis estadístico como cualquier otro programa (Archivo, Edición, etc.). Una vez que estamos seguros que no hay errores en la matriz, procedemos a realizar el análisis de la matriz, el análisis estadístico. En cada programa tales opciones varían, pero en cuestiones mínimas.

Ahora, comentaremos brevemente los programas más importantes y de dos de ellos señalaremos sus comandos generales.

Statistical Package for the Social Sciences SPSS® o PASW Statistics

El SPSS (Paquete Estadístico para las Ciencias Sociales) desarrollado en la Universidad de Chicago, es uno de los más difundidos. Contiene todos los análisis estadísticos que se describirán en este capítulo.

En América Latina, algunas instituciones educativas tienen versiones antiguas del SPSS; otras, versiones más recientes (PASW Statistics), ya sea en español o inglés. Existen versiones para Windows, Macintosh y UNIX. Desde luego, éstas sólo pueden utilizarse en computadoras con la capacidad necesaria para el paquete.

Como ocurre con todos los programas o software, SPSS® o PASW Statistics constantemente se actualiza con versiones nuevas en varios idiomas.¹ Asimismo, cada año surgen textos o manuales acordes con estas nuevas versiones. En el CD anexo el lector encontrará un manual que abarca las cuestiones esenciales de este paquete de análisis. Lo mejor para mantenerse al día en materia de SPSS/PASW es consultar su sitio en internet (www.spss.com/); o si éste llega a cambiar, con la palabra clave “SPSS” podemos encontrarlo en un directorio o mediante un motor de búsqueda como Google, Altavista, o cualquier otro. Para la actualización de manuales, las palabras claves serían: “SPSS manuals” (recorde mos que para cruzar palabras, éstas tienen que ir entre comillas “”).

Como ya se señaló, SPSS/PASW contiene las dos partes citadas que se denominan: *a*) vista de variables (para definiciones de las variables y consecuentemente, de los datos) y *b*) vista de los datos (matriz de datos).



¹ Actualmente SPSS cuenta con el programa básico (PASW) y múltiples derivaciones y aplicaciones. Por ejemplo: Quancept™ CATI (sistema para procesar y analizar entrevistas telefónicas), Amos™ (modelar ecuaciones estructurales) y el PASW Advanced Statistics (para estadística multivariada compleja). Para ver gran parte de las estadísticas avanzadas recomendamos: Sharpe, De Veaux y Velleman (2010), así como Madarassy (2010). El lector interesado puede descargar una versión de prueba de SPSS en el sitio www.spss.com

En ambas vistas se observan los comandos para operar en la parte superior. También, en la página de SPSS se puede “bajar” o “descargar” a la computadora una demostración del programa por un tiempo limitado.

El paquete SPSS/PASW trabaja de una manera muy sencilla: éste abre la matriz de datos y el investigador usuario selecciona las opciones más apropiadas para su análisis, tal como se hace en otros programas.

File (archivo): sirve para construir un nuevo archivo, localizar uno ya construido, guardar archivos, especificar impresora, imprimir, cerrar, enviar archivos por correo electrónico, entre otras funciones.

Edit (edición): se emplea para modificar archivos, manipular la matriz, buscar datos, copiar, cortar, eliminar y otras acciones de edición.

View (ver): como su nombre lo dice es para ver o visualizar la barra de estado, barra de herramientas, fuentes, cuadrícula (matriz), etiquetas y variables.

Data (datos): se insertan variables, sopesan casos, insertan casos, ordenan casos para limpiar archivos, fundir archivos (juntar varios archivos o matrices), segmentar archivos (por una variable o criterio; por ejemplo, la variable género, en este caso se realiza el análisis por submuestra segmentada, resultados para hombres y para mujeres), seleccionar casos, etcétera.

Transform (transformar): la función es de recodificar, conjuntar o unir y modificar variables y datos; categorizar variables; asignar rangos a casos, entre otras.

Analyze (analizar): se solicitan análisis estadísticos que básicamente serían:

1. Informes (resúmenes de casos, información de columnas y reglones).
2. Estadísticos descriptivos (tablas de frecuencias, medidas de tendencia central y dispersión, razones, tablas de contingencia).
3. Comparar medias (prueba *t* y análisis de varianza —ANOVA— unidireccional).
4. Modelo lineal general (independiente o factor y dependiente, con covariable).
5. ANOVA (análisis de varianza factorial en varias direcciones).
6. Correlaciones (bivariada —dos— y multivariadas —tres o más—) para cualquier nivel de medición de las variables.
7. Regresión (lineal, curvilínea y múltiple).
8. Clasificación (conglomerados y análisis discriminante).
9. Reducción de datos (análisis de factores).
10. Escalas (fiabilidad y escalamiento multidimensional).
11. Pruebas no paramétricas.
12. Respuestas múltiples (escalas).
13. Validación compleja.
14. Series de tiempos.
15. Ecuaciones estructurales y modelamiento matemático.

Add-ons (agregados): mediante esta función (que no está incluida en todas las versiones ni variantes) se tiene acceso a análisis complejos como redes neurales, identificación de casos inusuales y varias pruebas estadísticas avanzadas).

El diagrama Q-Q Se utiliza para verificar qué tanto la distribución de nuestras variables es “normal”.

Graphs (gráficos): con esta función se solicitan gráficos (histogramas, de sectores o pastel, diagramas de dispersión, Pareto, Q-Q —solicitar normalización de distribuciones—, P-P, curva COR, etcétera).

Utilities (utilidades o herramientas): se definen ambientes, conjuntos, información sobre variables, etcétera.

S-plus: es para la adquisición, edición y transformación de datos, la línea de comandos, métodos estadísticos básicos con S-Plus y R, gráficos estadísticos básicos con S-Plus y R, métodos estadísticos multivariados avanzados y creación de funciones propias con S-Plus.

Window (ventana): sirve para moverse a través de archivos y hacia otros programas.

Help (ayuda): cuenta con contenidos de ayuda, cómo utilizar SPSS, comandos, guías, “asesor estadístico” y demás elementos aplicados al paquete (con índice).

Minitab®

Es un paquete que goza de popularidad por su relativamente bajo costo. Incluye un considerable número de pruebas estadísticas, y cuenta con un tutorial para aprender a utilizarlo y practicar; además, es muy sencillo de manejar.

Minitab tiene un sitio web (<http://www.minitab.com/>) en la cual podemos acceder a un demo gratuito del programa por tiempo limitado.

Para comenzar a utilizar Minitab, se abre una sesión (la cual es definida con nombre y fecha), y se abre una matriz u hoja de trabajo (*worksheet*) (en la parte superior de la pantalla aparece la sesión y en la parte inferior se presenta la matriz). Se definen las variables (*C* —columnas—): nombre, formato (numérico, texto, fecha/tiempo), ancho (en dígitos), su descripción y orden de los valores. Los renglones o filas son casos. Los análisis realizados aparecen en la sesión (parte o pantalla superior) y las gráficas se reproducen en recuadros.

Sus comandos incluyen:

File (archivo): para construir un nuevo archivo, localizar uno ya construido, guardar o abrir archivos, abrir una gráfica de Minitab, especificar impresora, imprimir, cerrar, entre otras funciones.

Edit (edición): útil para modificar archivos, buscar datos, copiar, cortar y eliminar celdas, conectar Minitab con otras aplicaciones, etcétera.

Data (datos): funciones para asignar códigos a columnas, dividir la matriz, copiar columnas, eliminar columnas y renglones o filas, establecer rangos, recodificar, cambiar el tipo de datos, desplegar datos, mostrar los datos de la hoja de trabajo en la ventana de sesión, entre otros.

Calc (calcular): calcula las estadísticas de columnas y filas, distribuciones de probabilidad, matrices, estandarizaciones, operaciones aritméticas.

Stat (estadísticas): de manera fundamental, ejecuta los siguientes tipos de estadísticas:

1. Básicas: descriptivas, correlación, covarianza, *chi*-cuadrada, prueba *t*, prueba de hipótesis acerca de la media poblacional...
2. Regresión lineal y múltiple.
3. Análisis de varianza (ANOVA) unidireccional y factorial.
4. DOE (análisis para diseños experimentales, análisis de respuestas).
5. Diagramas (control charts) (de atributos, multivariados, de tiempo) individuales y grupales.
6. Diagramas de dispersión, Pareto, causa-efecto...
7. Confiabilidad.
8. Análisis multivariado: análisis de factores (validación), análisis discriminante, análisis de conglomerados, de correspondencia simple o múltiple.
9. Series de tiempos: autocorrelación, correlación parcial, correlación cruzada, entre otras.
10. Tablas: tabulación cruzada, *chi*-cuadrada.
11. Estadística no paramétrica.
12. EDA (análisis exploratorio de datos, diagramas de caja, fotograma, etcétera).
13. Poder y tamaño de muestra (1-muestra *z*, 1-muestra-*t*, 2-muestra-*t*, ANOVA y otras. Sirve para determinar si el tamaño de muestra es apropiado para varias pruebas estadísticas).

Graph (gráfica): solicitar gráficos (histogramas, barras de pastel, diagramas de dispersión, Pareto, series de tiempos, etcétera).

Editor (editor): mover columnas, redefinir columnas, insertar columnas, buscar, ir a un caso, entre otras acciones.

Tools (herramientas): definir ambientes, conjuntos, información sobre variables, conexión a internet, consultas, etcétera.

Window (ventana): sirve para moverse a través de archivos y hacia otros programas, minimizar ventanas y demás funciones similares en otros programas.

Help (ayuda): cuenta con contenidos de ayuda, cómo utilizar Minitab, comandos, guías y demás elementos de Windows aplicados al paquete. En la figura 10.2 se muestra una vista de la pantalla de Minitab.

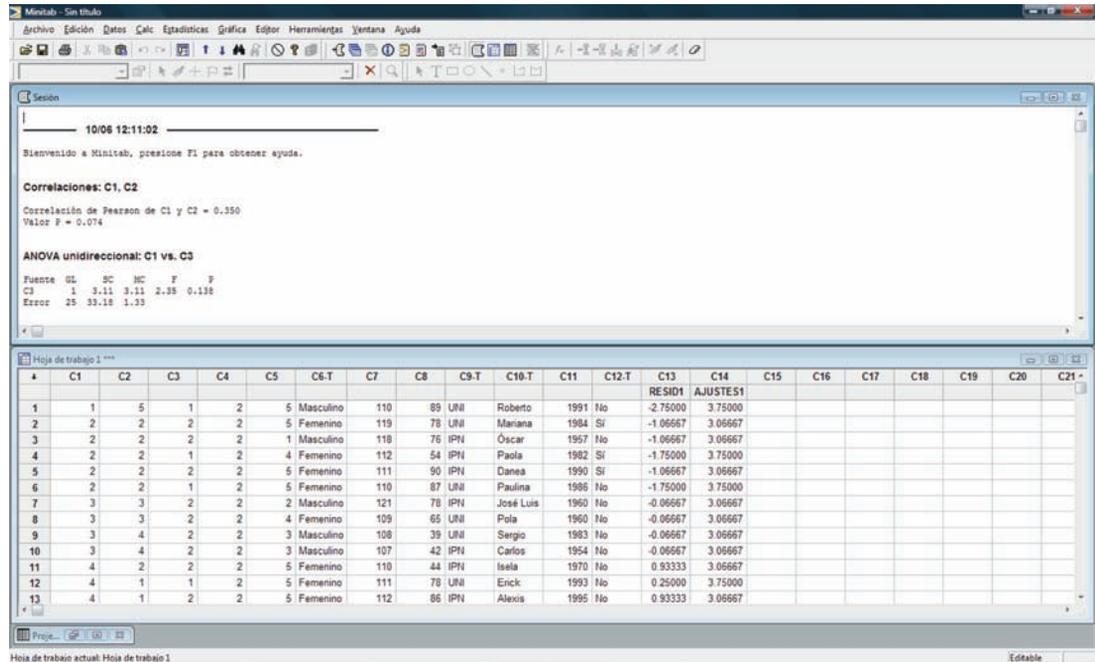


Figura 10.2 Pantalla de Minitab.

Otro programa de análisis sumamente difundido es el SAS (Sistema de Análisis Estadístico), que fue diseñado en la Universidad de Carolina del Norte. Es muy poderoso y su utilización se ha incrementado notablemente. Es un paquete muy completo para computadoras personales que contiene una variedad considerable de pruebas estadísticas.

En el CD se incluye un programa (software) sencillo que hemos titulado STATS, con los análisis bivariados más elementales para comenzar a practicar y comprender las pruebas básicas. Asimismo, en internet existen diversos programas gratuitos de análisis estadístico para cualquier ciencia o disciplina.

Por lo general se elige el programa de análisis que está disponible en nuestra institución educativa, centro de investigación u organización de trabajo, o el que podamos comprar u obtener en internet. Todos los programas mencionados son excelentes opciones. Cualquiera nos sirve, solamente que debemos seleccionar uno. Recomendamos que en el centro de cómputo de su institución soliciten información respecto de los programas disponibles.

Paso 2: ejecutar el programa

En el caso de SPSS y Minitab, ambos paquetes son fáciles de usar, pues lo único que hay que hacer es solicitar los análisis requeridos seleccionando las opciones apropiadas. Obviamente antes de tales análisis, se debe verificar que el programa “corra” o funcione en nuestra computadora. Comprobado esto, comienza la ejecución del programa y la tarea analítica.

Paso 3: explorar los datos

En esta etapa, inmediata a la ejecución del programa, se inicia el análisis. Cabe señalar que si hemos llevado a cabo la investigación reflexionando paso a paso, esta etapa es relativamente sencilla, porque: 1) formulamos la pregunta de investigación que pretendemos contestar, 2) visualizamos un alcance (exploratorio, descriptivo, correlacional y/o explicativo), 3) establecimos nuestras hipótesis (o estamos

conscientes de que no las tenemos), 4) definimos las variables, 5) elaboramos un instrumento (conocemos qué ítems miden qué variables y qué nivel de medición tiene cada variable: nominal, ordinal, de intervalos o razón) y 6) recolectamos los datos. Sabemos qué deseamos hacer, es decir, tenemos claridad.

La exploración típica se muestra en la figura 10.3 (que se ilustra utilizando el programa SPSS, ya que, insistimos, ésta puede variar de programa en programa en cuanto a comandos o instrucciones, pero no en lo referente a las funciones implementadas). Algunos conceptos pueden, por ahora, no significar algo para el lector que se inicia en los menesteres de la investigación, pero éstos se irán explicando a lo largo del capítulo.

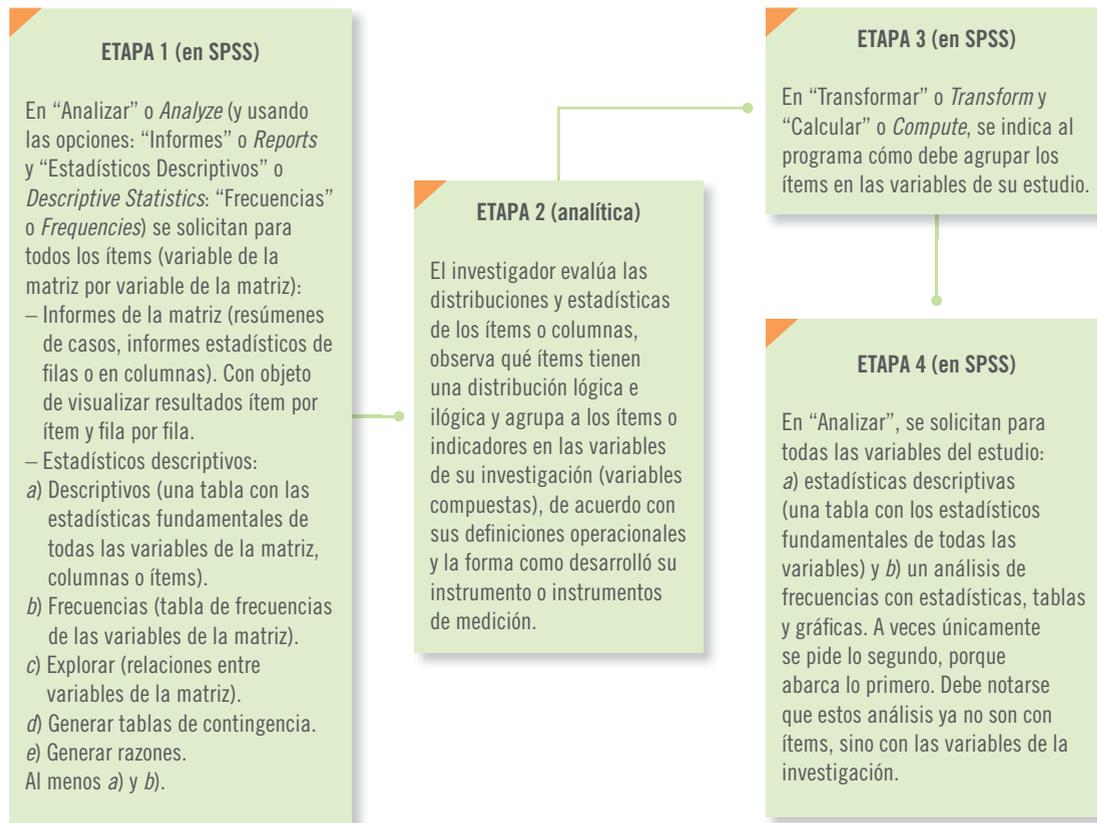


Figura 10.3 Secuencia más común para explorar datos en SPSS.

Veamos ahora los conceptos estadísticos que se aplican a la exploración de datos, pero antes de proseguir es necesario realizar un par de apuntes, uno sobre las *variables del estudio* y las *variables de la matriz de datos*, y el otro sobre los factores de los que depende el análisis.



Apunte 1

Desde el final del capítulo anterior, se introdujo el concepto de *variable de la matriz de datos*, que es distinto del concepto *variable de la investigación*. Las **variables de la matriz de datos** son columnas o ítems. Las **variables de la investigación** son las propiedades medidas y que forman parte de las hipótesis o que se pretenden describir (género, edad, actitud hacia el presidente municipal, inteligencia, duración de un material, etc.). En ocasiones, las variables de la investigación requieren un único ítem o indicador para ser medidas (como en la tabla 10.2 con la variable “tipo de escuela a la que asiste”), pero en otras se necesitan varios ítems para tal finalidad. Cuando sólo se precisa de un ítem o

Variables de la matriz de datos
 Son columnas o ítems.

Variables de la investigación Son las propiedades medidas y que forman parte de las hipótesis o que se pretenden describir.

indicador, las variables de la investigación ocupan una columna de la matriz (una variable de la matriz). Pero si están compuestas de varios ítems, ocuparán tantas columnas como ítems (o variables en la matriz) las conformen. Esto se ejemplifica en la tabla 10.2 con los casos de la variable “satisfacción respecto al superior” y “moral de los empleados”.

▲ **Tabla 10.2** Ejemplos de variables de investigación y formulación de ítems

Variable: tipo de escuela a la que asiste (con un ítem)	Variable: satisfacción respecto al superior (con tres ítems)	Variable: moral de los empleados (con cinco ítems)
<p>¿Asiste a una escuela pública o privada?</p> <p>1 Escuela pública 2 Escuela privada</p>	<p>1. ¿En qué medida está usted satisfecho con su superior inmediato?</p> <p>1 Sumamente insatisfecho 2 Más bien insatisfecho 3 Ni insatisfecho ni satisfecho 4 Más bien satisfecho 5 Sumamente satisfecho</p>	<p>1. “En el departamento donde trabajo nos mantenemos unidos”</p> <p>5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo</p>
	<p>2. ¿Qué tan satisfecho está usted con el trato que recibe de parte de su superior inmediato?</p> <p>1 Sumamente insatisfecho 2 Más bien insatisfecho 3 Ni insatisfecho ni satisfecho 4 Más bien satisfecho 5 Sumamente satisfecho</p>	<p>2. “La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros”.</p> <p>5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo</p>
	<p>3. ¿Qué tan satisfecho está con la orientación que le proporciona su superior inmediato para que usted realice su trabajo?</p> <p>1 Sumamente insatisfecho 2 Más bien insatisfecho 3 Ni insatisfecho ni satisfecho 4 Más bien satisfecho 5 Sumamente satisfecho</p>	<p>3. “En mi departamento nos mantenemos en contacto permanentemente”.</p> <p>5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo</p>
		<p>4. “En mi departamento nos reunimos con frecuencia para hablar tanto de asuntos de trabajo como de cuestiones personales”.</p> <p>5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo</p>
		<p>5. “En mi trabajo todos nos llevamos bien”.</p> <p>5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo</p>
<p>Esta variable es medida por una sola pregunta y ocupa una columna o variable de la matriz.</p>	<p>Esta variable es medida por tres preguntas y ocupa tres columnas o variables de la matriz.</p>	<p>Esta variable es medida por cinco preguntas y ocupa cinco columnas o variables de la matriz.</p>

Y cuando las variables de la investigación se integran de varios ítems o variables en la matriz, las columnas pueden ser continuas o no (estar ubicadas de manera seguida o en distintas partes de la matriz). En el tercer ejemplo (variable “moral de los empleados”), las preguntas podrían ser las núme-

ros: 1, 2, 3, 4 y 5 del cuestionario, entonces las primeras cinco columnas de la matriz representarán a estos ítems. Pero pueden ubicarse en distintos segmentos del cuestionario (por ejemplo, ser las preguntas 1, 5, 17, 22 y 38), entonces las columnas que las representen se ubicarán de forma discontinua (serán las columnas o variables de la matriz 1, 5, 17, 22 y 38), porque regularmente la secuencia de las columnas corresponde a la secuencia de los ítems en el instrumento de medición.

Esta explicación la hacemos porque hemos visto que varios estudiantes confunden las variables de la matriz de datos con las variables del estudio. Son cuestiones vinculadas pero distintas.

Cuando una variable de la investigación está integrada por diversas variables de la matriz o ítems suele denominarse *variable compuesta* y su puntuación total es el resultado de adicionar los valores de los reactivos que la conforman. Tal vez el caso más claro lo es la escala Likert, donde se suman las puntuaciones de cada ítem y se logra la calificación final. A veces la adición es una sumatoria, otras ocasiones es multiplicativa o de otras formas, según se haya desarrollado el instrumento. Al ejecutar el programa y durante la fase exploratoria, se toma en cuenta a *todas las variables de la investigación* e ítems y se considera a las *variables compuestas*, entonces se indica en el programa cómo están constituidas, mediante algunas instrucciones (en cada programa son distintas en cuanto al nombre, pero su función es similar). Por ejemplo, en SPSS se crean nuevas variables compuestas en la matriz de datos con el comando “Transformar” y luego con el comando “Calcular” o “Computar”, de este modo, se construye la variable compuesta mediante una expresión numérica. Revisemos un ejemplo.

En el caso de la variable “moral en el departamento de trabajo”, podríamos asignar las siguientes columnas (en el supuesto de que fueran continuas) a los cinco ítems, tal como se muestra en la tabla 10.3.

Y tener la siguiente matriz:

EJEMPLO					
	fr1	fr2	fr3	fr4	fr5
1	1	2	2	4	3
2	2	2	2	2	2
K	2	3	2	2	3

▲ **Tabla 10.3** Ejemplo con la variable moral

Variable de la investigación: moral	Variable de la matriz que corresponde a la variable de la investigación	Ubicación en la matriz
1. “En el departamento donde trabajo nos mantenemos unidos”. 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 1 (fr1)	Columna 1
2. “La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros”. 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 2 (fr2)	Columna 2
		(continúa)



▲ **Tabla 10.3** Ejemplo con la variable moral (*continuación*)

Variable de la investigación: moral	Variable de la matriz que corresponde a la variable de la investigación	Ubicación en la matriz
3. “En mi departamento nos mantenemos en contacto permanentemente”. 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 3 (fr3)	Columna 3
4. “En mi departamento nos reunimos con frecuencia para hablar tanto de asuntos de trabajo como de cuestiones personales”. 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 4 (fr4)	Columna 4
5. “En mi trabajo todos nos llevamos muy bien”. 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 5 (fr5)	Columna 5

En las opciones “Transformar” y “Calcular” o “Computar” el programa nos pide que indiquemos el nombre de la nueva variable (en este caso la compuesta por cinco frases): *moral*. Y nos solicita que desarrollemos la expresión numérica que corresponda a esta variable compuesta: $fr1+fr2+fr3+fr4+fr5$ (automáticamente el programa realiza la operación y agrega la nueva variable compuesta “moral” a la matriz de datos y realiza los cálculos, y ahora sí, la *variable del estudio* es una variable más de la matriz de datos). La matriz se modificaría de la siguiente manera:



EJEMPLO						
	fr1	fr2	fr3	fr4	fr5	Moral
1	1	2	2	4	3	12
2	2	2	2	2	2	10
K	2	3	2	2	3	12

Desde luego, para mantener esta variable debemos demostrar que fue medida de forma confiable y válida, así como evaluar si todos los ítems aportan favorablemente a ambos elementos o algunos no. Y en lugar de una suma, la variable *moral* podría ser un promedio de las cinco frases o variables de la matriz (como ya se mencionó en el tema de la escala Likert). Entonces, la expresión en “Calcular” hubiera sido: $(fr1+fr2+fr3+fr4+fr5)/5$, y los valores en “moral” serían:



EJEMPLO						
	fr1	fr2	fr3	fr4	fr5	Moral
1	1	2	2	4	3	2.4
2	2	2	2	2	2	2.0
K	2	3	2	2	3	2.4

Por último, las variables de la investigación son las que nos interesan, ya sea que estén compuestas por uno, dos, diez, 50 o más ítems. El primer análisis es sobre los ítems, únicamente para explorar; el análisis descriptivo final es sobre las *variables del estudio*.

Apunte 2

Los análisis de los datos dependen de tres factores:

- a) El *nivel de medición* de las variables.
- b) La manera como se hayan formulado las *hipótesis*.
- c) El *interés del investigador*.

Por ejemplo, los análisis que se aplican a una variable nominal son distintos a los de una variable por intervalos. Se sugiere recordar los niveles de medición vistos en el capítulo anterior.

El investigador busca, en primer término, describir sus datos y posteriormente efectuar análisis estadísticos para relacionar sus variables. Es decir, realiza análisis de estadística descriptiva para cada una de las variables de la matriz (ítems) y luego para cada una de las variables del estudio, finalmente aplica cálculos estadísticos para probar sus hipótesis. Los tipos o métodos de análisis cuantitativo o estadístico son variados y se comentarán a continuación; pero cabe señalar que el análisis no es indiscriminado, cada método tiene su razón de ser y un propósito específico; por ello, no deben hacerse más análisis de los necesarios. La estadística no es un fin en sí misma, sino una herramienta para evaluar los datos.

Estadística descriptiva para cada variable

La primera tarea es describir los datos, los valores o las puntuaciones obtenidas para cada variable. Por ejemplo, si aplicamos a 2 112 niños el cuestionario sobre los usos y las gratificaciones que la televisión tiene para ellos, ¿cómo pueden describirse estos datos? Esto se logra al describir la distribución de las puntuaciones o frecuencias de cada variable.

¿Qué es una distribución de frecuencias?

OQ2

Una **distribución de frecuencias** es un conjunto de puntuaciones ordenadas en sus respectivas categorías y generalmente se presenta como una tabla.

La tabla 10.4 muestra un ejemplo de una distribución de frecuencias.

Distribución de frecuencias Conjunto de puntuaciones ordenadas en sus respectivas categorías.

EJEMPLO

En un estudio entre 200 personas latinas que viven en el estado de California, Estados Unidos,² se les preguntó: ¿cómo prefiere que se refieran a usted en cuanto a su origen étnico? Las respuestas fueron:

▲ **Tabla 10.4** Ejemplo de una distribución de frecuencias

Variable: preferencias al referir el origen étnico (nombrada en SPSS: prefoe)		
Categorías	Códigos (valores)	Frecuencias
Hispano	1	52
Latino	2	88
Latinoamericano	3	6
Americano	4	22
Otros	5	20
No respondieron	6	12
Total		200

² Encuesta con 7% de margen de error (University of Southern California y Bendixen & Associates, 2002).

A veces, las *categorías* de las distribuciones de frecuencias son tantas que es necesario resumirlas. Por ejemplo, examinaremos detenidamente la distribución de la tabla 10.5. Esta distribución podría compendiarse como en la tabla 10.6.

▲ **Tabla 10.5** Ejemplo de una distribución que necesita resumirse

Variable: calificación en la prueba de motivación	
Categorías	Frecuencias
48	1
55	2
56	3
57	5
58	7
60	1
61	1
62	2
63	3
64	2
65	1
66	1
68	1
69	1
73	2
74	1
75	4
76	3
78	2
80	4
82	2
83	1
84	1
86	5
87	2
89	1
90	3
92	1
TOTAL	63

▲ **Tabla 10.6** Ejemplo de una distribución resumida

Variable: calificación en la prueba de motivación	
Categorías	Frecuencias
55 o menos	3
56-60	16
61-65	9
66-70	3
71-75	7
76-80	9
81-85	4
86-90	11
91-96	1
TOTAL	63