

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 lo 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.

Volúmenes

En este caso supongamos que un estudioso del periodismo quiere hacer un análisis de contenido de los editoriales de los tres principales diarios de la ciudad durante los periodos del porfiriato en México, el gobierno sandinista en Nicaragua o el franquismo en España. El investigador va a la Hemeroteca Nacional y encuentra —por ejemplo— que los diarios son encuadernados por trimestre y año, lo cual le proporciona un marco de referencia ideal, a partir de donde seleccionará n volúmenes para su análisis. Supongamos, en el caso mexicano, que encuentra que el volumen X , que contiene el periódico *El Hijo del Ahuizote* (enero-marzo, 1899), falta en la hemeroteca. ¿Qué hace? Pues redefine la población, manifestando explícitamente que de N volúmenes tiene 99% de los elementos y, a partir de este nuevo número de N , calcula su muestra n y la selecciona.

Periodos registrados

En la investigación que se vincula con sesiones terapéuticas, entrevistas con pacientes o reos, etc., se dispone de los listados de archivos que por lo común guardan las instituciones.

El investigador debe buscar posibles listados o bases de datos en cualquier parte para que pueda obtener el tamaño de su población. Por ejemplo, para analizar anuncios en los medios electrónicos de comunicación, hay empresas que tienen un registro de todos los anuncios transmitidos; la hora exacta y el canal o frecuencia radial en que fueron difundidos, su duración, la versión del anuncio de que se trata, etc. El ingenio es el propio límite para ello.

Tamaño óptimo de una muestra

Las muestras probabilísticas requieren dos procedimientos básicos: 1) la determinación del tamaño de la muestra y 2) la selección aleatoria de los elementos muestrales. El primer procedimiento fue descrito en su modalidad más simple en la sección sobre el tamaño de la muestra. Precisar adecuadamente el tamaño de la muestra puede tornarse muy 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 (8.6, 8.7 y 8.8), 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
	1737

La muestra de los barómetros de opinión en España es nacional,⁷ incluye personas de ambos géneros, de 18 años o más y su tamaño es alrededor de 2 500 casos (Berganza y García, 2005). 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).

⁷ Los barómetros son *surveys* o encuestas de alcance nacional o continental e incluyen cuestiones políticas, económicas, sociales y de actualidad.

En cambio, el barómetro del Real Instituto Elcano (BRIE) en España comprende a 1 200 individuos (Berganza y García, 2005).

El eurobarómetro es otra encuesta que abarca a diversos países de la Unión Europea 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).

En la tabla 8.6⁸ 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.

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 8.7 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 sujetos 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.

Otra tabla que nos ayuda a comprender el tema que estamos analizando es la 8.9, 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.

Las tablas 8.6 a 8.8 se construyeron con base en artículos de investigación publicados en revistas especializadas 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; Hernández, Fernández y Baptista, 2003).

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

Tabla 8.6 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 Las tablas 8.6 y 8.7 fueron adaptadas de Sudman (1976, pp. 86-87).

Tabla 8.7 Muestras típicas de estudios sobre poblaciones humanas y organizaciones

Número de subgrupos	Población de sujetos u hogares		Poblaciones de organizaciones	
	Nacionales	Regionales	Nacionales	Regionales
Ninguno-pocos	1 000-1 500	200-500	200-500	50-200
Promedio	1 500-2 500	500-1 000	500-1 000	200-500
Muchos	2 500 +	1 000 +	1 000 +	500 +

Tabla 8.8 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.

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, sobre los valores de una población. 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 (ver figura 8.5).

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.



Figura 8.5 Distribución normal con muestras de 100 o más unidades muestrales.

... ¿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 muchas investigaciones cuantitativas y cualitativas. No las revisaremos ahora, sino en el capítulo sobre muestras cualitativas. Por el momento comentaremos su esencia y utilidad desde una perspectiva cuantitativa y ejemplificaremos la diferencia con las muestras probabilísticas.

La *muestra dirigida* selecciona sujetos "típicos" con la vaga esperanza de que sean casos 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 en este sentido implican muchas 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 dirigidas tienen un valor limitado a la muestra en sí, mas no a la población. Es decir, los datos no pueden generalizarse a ésta, la cual no se consideró en sus parámetros ni en sus elementos para obtener la muestra. Recordemos que, en las muestras de este tipo, la elección de los sujetos 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 sujetos 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 —si se procede cuidadosamente y con una profunda inmersión inicial en el campo— 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.

En el siguiente ejemplo compararemos, con términos muy sencillos, una muestra probabilística simple con una dirigida.

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 mil alumnos y alumnas. Si obtuviera una muestra aleatoria simple, su procedimiento sería el que se muestra en la figura 8.6.

Pero, si en lugar de este procedimiento, la investigadora determinara que requiere 100 alumnos y seleccionara a los cien primeros que entraran a la cafetería, su muestra sería no probabilística; o bien, si eligiera nada más a los representantes de los grupos (designados por la mesa directiva de alumnos o las autoridades de la escuela), también tendría una muestra de este tipo. En el primer caso, no todos los estudiantes tienen la misma probabilidad de ser seleccionados al inicio, porque no todos asisten a la cafetería o no todos acuden ese día; y en el segundo, no todos son representantes (quizás éstos sean los más populares o aplicados de cada grupo y poseen una cualidad que está presente sólo en ciertos casos).

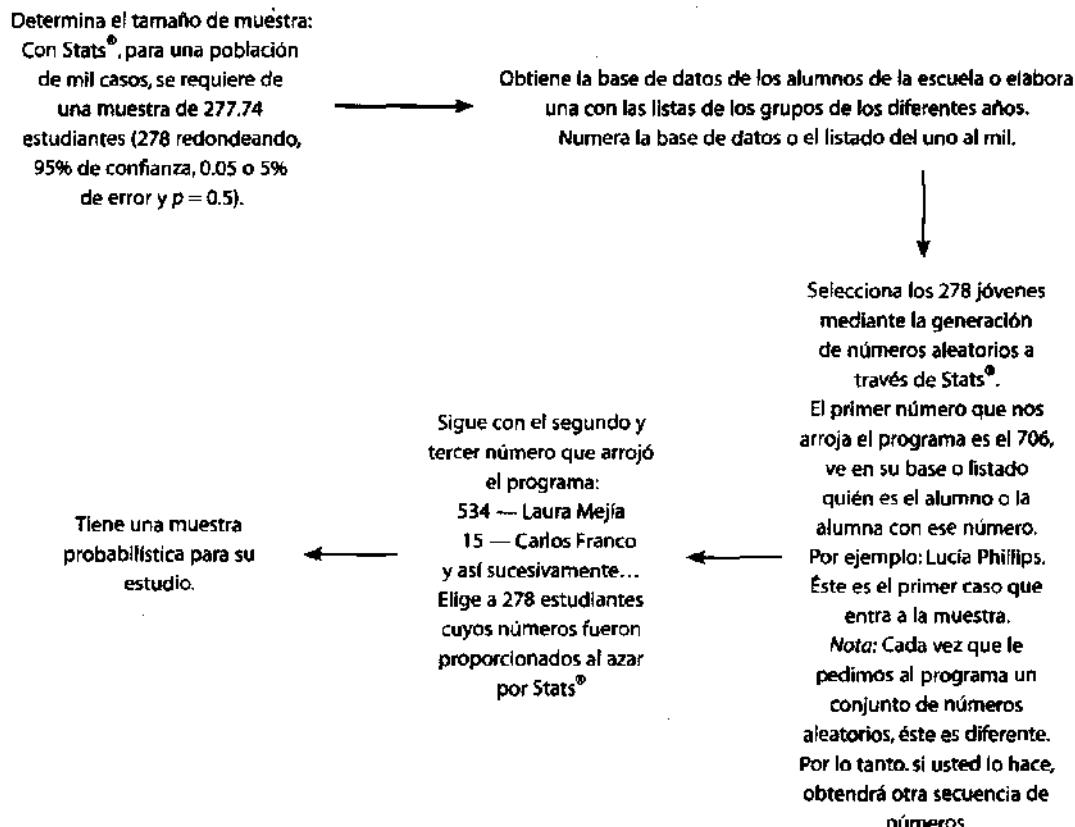


Figura 8.6 Ejemplo del procedimiento para una muestra aleatoria simple.

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 tres dígitos del número telefónico). Luego, los demás dígitos del número a marcar pueden ser generados al azar de acuerdo a los casos que requerimos para la muestra (n). Es posible identificar qué intercambios son usados de forma primaria para teléfonos residenciales y enfocar el muestreo en ese subgrupo. Para mayores referencias de esta técnica recomendamos Fowler (2002).

Muestra multietapas o polietápica

Este concepto significa que para extraer la muestra hemos utilizado diversos procedimientos. Por ejemplo, en la investigación sobre la televisión y los niños de la ciudad de México, el procedi-

miento de selección implicó estratos y racimos: en una primera etapa se seleccionarían “escuelas” y en una segunda “niños”.

En un estudio nacional, primero podrían elegirse ciudades, luego en cada ciudad, barrios; posteriormente en cada barrio, calles; luego casas y al final personas. En cada etapa el procedimiento es aleatorio.

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 o no probabilísticas, 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.



- 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 la definición de 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 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 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 deli-

rado 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 dos etapas, ambas con procedimientos probabilísticos. En la primera se seleccionan los racimos: escuelas, organizaciones, salones de clase; en la segunda 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 tres procedimientos de selección: 1) tómbola, 2) cuadro de números aleatorios o el uso del subprograma de números aleatorios del Stats® y 3) 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 recurren 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 sujetos u objetos de estudio 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
Listados
Muestra
Muestra probabilística
Muestra no probabilística o dirigida
Población

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



EJERCICIOS

1. 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.

2. 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 significativas existen entre los comerciales de la televisión española, la argentina y la venezolana?

3. Seleccione dos estudios de alguna publicación científica (véase apéndice 1 incluido en el CD) y/o dos tesis. Analice los siguientes aspectos: a) ¿Cuál es el problema de investigación? b) ¿Cuál es la muestra? c) ¿Cómo fue elegida? d) ¿Son adecuadas la muestra y el procedimiento de muestreo para el problema que se investigó? e) ¿Cuáles son los principales resultados o conclusiones? f) ¿Dichos resultados son generalizables a una población mayor? g) 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.
4. 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.	

CLIENTE	NECESIDAD	TIPO DE MUESTRA
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 cosmetología	¿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?	

5. 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 5 000 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 5 000 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 apéndice 3 del CD anexo.

LA TELEVISIÓN Y EL NIÑO

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

1) Límites de población:

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

2) Proceso de selección:

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

Se excluiría a escuelas del turno vespertino y las diseñadas para niños con capacidades diferentes o habilidades especiales. La selección también estratificaría 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 lo tanto, se elegirían 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 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. Posteriormente, de cada escuela se elegirían los niños para conformar la muestra final.

Una vez hechos los cálculos, se determinó que de cada estrato se seleccionarían cuatro escuelas (es decir n es igual a 32 escuelas ubicadas en diversas colonias que incluirán a todas las delegaciones (municipalidades). En la segunda etapa se seleccionarían por muestreo aleatorio simple los niños de cada escuela. En el ejemplo, 264 infantes por escuela de 4°, 5° y 6° grados (88 por cada uno). Una muestra total de 2 112 que podría implicar ajustes y reemplazos.

EL CLIMA ORGANIZACIONAL

La muestra es dirigida o no probabilística, se seleccionarán organizaciones de dife-

rentes características para comenzar a validar el instrumento en el medio laboral mexicano: a) Una empresa transnacional de la industria de la transformación con 500 o más empleados, b) una empresa nacional de la industria de la transformación con más de 500 empleados, c) una institución de educación superior con más de cinco mil personas empleadas, d) dos instituciones de educación media y superior con más de 200 personas trabajando en ella y e) cinco empresas comerciales y de servicio con 100 a 200 empleados. De todas ellas se hará un censo (es decir, se pretenderá abarcar a todos los empleados), salvo el caso de la institución educativa con cinco mil personas (de la cual se extraerá una muestra aleatoria simple en donde n es igual a 357 casos, con 5% de error).

El propósito no es generalizar la validación a todos los individuos que trabajan en el país, sino comenzar a probar la solidez del instrumento para medir el clima organizacional. Además debemos recordar que los resultados del clima laboral de una empresa son temporales y únicos de ésta. No existen dos organizaciones iguales (simplemente, la historia de cada una es distinta).

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,

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 evaluarán seis grupos escolares que serán 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.

además de procurar la comunicación de los resultados.

Álvaro Camacho Medina

Docente

Facultad de Mercadeo y Publicidad

Politécnico Gran Colombiano

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.

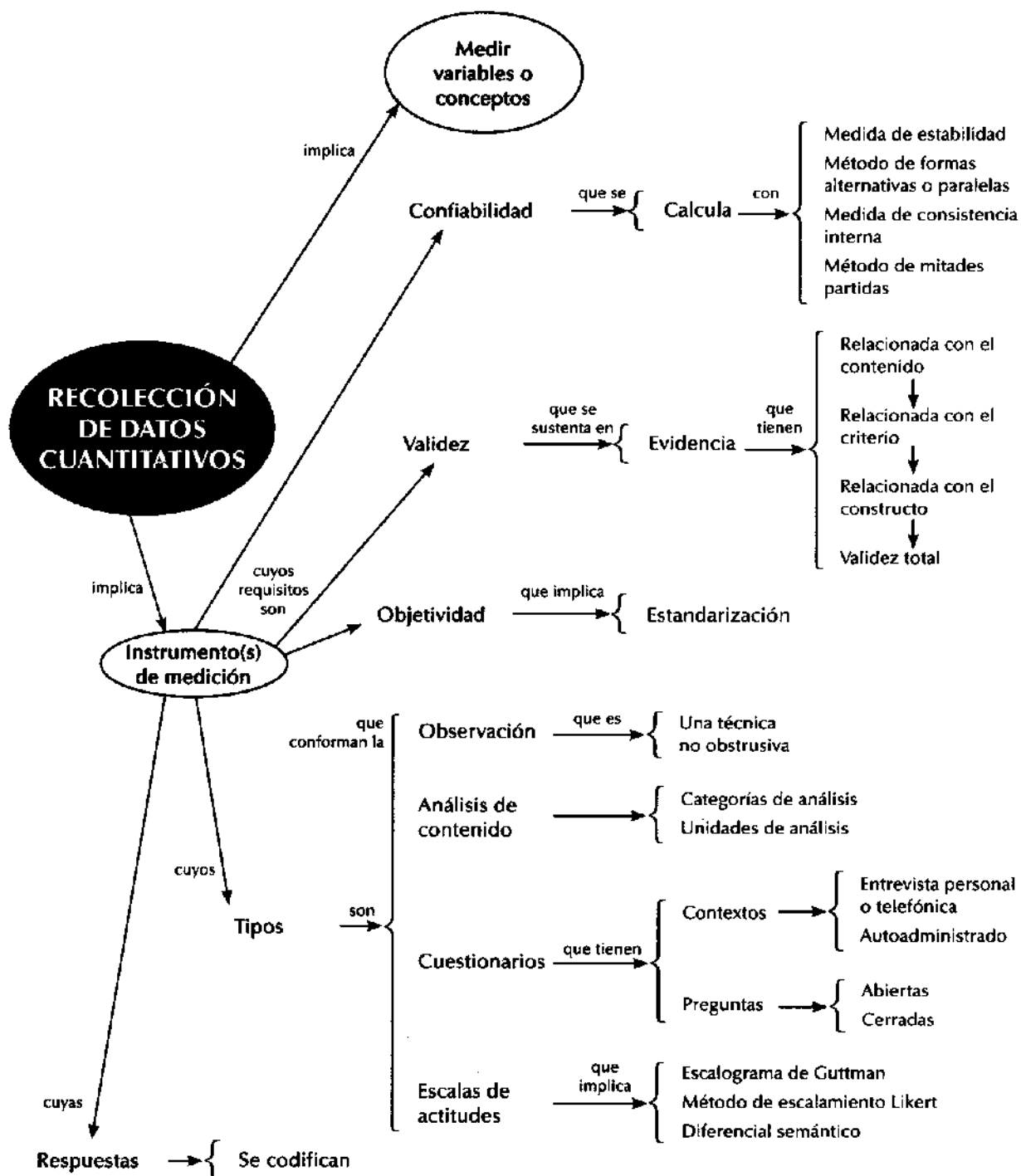
Ing. Guillermo Evangelista Benites

Docente principal

Facultad de Ingeniería Química

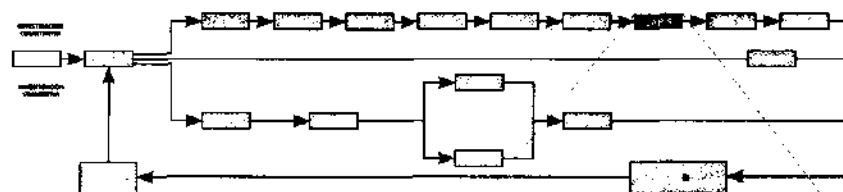
Universidad Nacional de Trujillo

Trujillo, Perú



Capítulo 9

Recolección de los datos cuantitativos



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.

OBJETIVOS DE APRENDIZAJE

Al terminar este capítulo, el alumno será capaz de:

- Visualizar diferentes métodos para recolectar datos cuantitativos.
- Comprender el significado de "medir" y su importancia en el proceso cuantitativo.
- Comprender los requisitos que toda recolección de datos debe incluir.
- Conocer los principales instrumentos para recolectar datos cuantitativos.
- Elaborar y aplicar los diferentes instrumentos de recolección de datos cuantitativos.
- 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: entrevistas y cuestionarios, escalas de actitudes, formatos para evaluar contenidos de comunicaciones, sistemas de observación, pruebas estandarizadas y otras más. Por último, se examina el procedimiento de codificación de datos cuantitativos y la forma de prepararlos para su análisis.

... ¿Qué implica la etapa de recolección de los datos?

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, cualidades o variables de los participantes, casos, sucesos, comunidades u objetos involucrados en la investigación.

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 vamos a obtener 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 (contenidas 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 y no olvidemos que todos los atributos, cualidades y variables deben ser medibles.

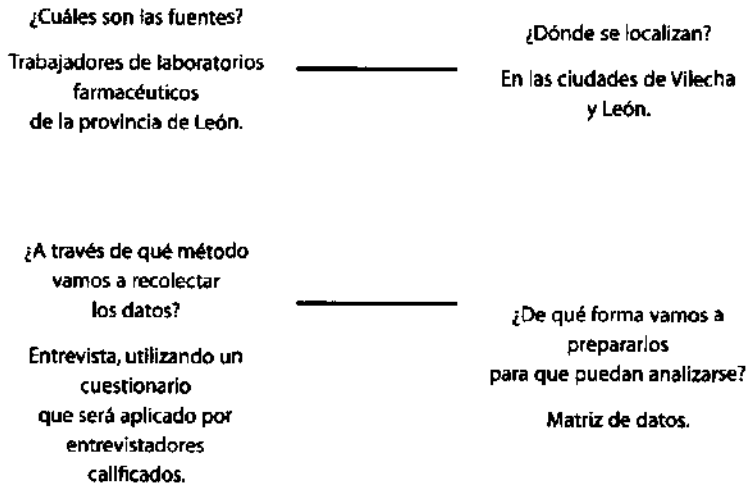
Con la finalidad de **recolectar datos** disponemos de una gran variedad de instrumentos o técnicas, tanto cuantitativas como cualitativas, es por ello que en un mismo estudio podemos utilizar ambos tipos. Incluso, hay instrumentos como la prueba de propósito vital (PIL) (que evalúa el propósito de vida de una persona) de Crumbaugh y Maholick (1969) que, en la misma prueba, tienen una parte cuantitativa y una cualitativa (Brown, Ashcroft y Miller, 1998).

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

Planteamiento

Objetivo: Analizar la relación entre la satisfacción de las recompensas, el grado de responsabilidad, el desarrollo personal y la confianza en sí mismo en los trabajadores de la industria farmacéutica.

Pregunta: ¿Cuál es la relación entre la satisfacción de las recompensas, el grado de responsabilidad, el desarrollo personal y la confianza en sí mismo en los trabajadores de la industria farmacéutica?

PLAN

1. Las variables a medir son: satisfacción de las recompensas, grado de responsabilidad, desarrollo personal, confianza en sí mismo.
2. Las definiciones operacionales: Encuesta del clima organizacional (Training House, 2000) que mide las variables de interés.
3. La muestra: 300 trabajadores.
4. Recursos disponibles: económicos, suficientes. Tiempo: un mes.

Figura 9.1 Ejemplo de plan para la obtención de datos.

• • ¿Qué significa medir?

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 tráfico 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).

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 alienación, 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.

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 (Carmines y Zeller, 1991). Así, los registros del instrumento de medición representan valores visibles de conceptos abstractos. Un **instrumento de medición** adecuado es aquel que registra datos observables que representan verdaderamente los conceptos o las variables que el investigador tiene en mente. 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

Medición Proceso que vincula conceptos abstractos con indicadores empíricos.

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

las variables que tenemos en mente. Si no es así, nuestra medición es deficiente; por lo 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).

¿Qué requisitos debe cubrir un instrumento de medición?

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

La confiabilidad

La **confiabilidad** de un instrumento de medición se refiere al grado en que su aplicación repetida al mismo sujeto 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 (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 95 y 150). Los resultados no son coherentes, pues no se puede “confiar” en ellos.

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.

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

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. Una prueba sobre conocimientos de historia tiene que medir esto y no conocimientos de literatura histórica. Un método para medir el rendimiento bursátil tiene que medir precisamente esto y no la imagen de una empresa. En apariencia es sencillo lograr

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

Primera aplicación		Segunda aplicación		Tercera aplicación	
Martha	130	Laura	131	Luis	140
Laura	125	Luis	130	Teresa	129
Gabriel	118	Marco	127	Martha	124
Luis	112	Gabriel	120	Rosa María	120
Marco	110	Chester	118	Laura	109
Rosa María	110	Teresa	118	Chester	108
Chester	108	Martha	115	Gabriel	103
Teresa	107	Rosa María	107	Marco	101

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.

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

La **validez** es un concepto del cual pueden tenerse diferentes tipos de evidencia (Bostwick y Kyte, 2005; Creswell, 2005; Wiersma y Jurs, 2005; y Gronlund, 1990): 1) *evidencia relacionada con el contenido*, 2) *evidencia relacionada con el criterio* y 3) *evidencia relacionada con el constructo*. Hablaremos de cada una de ellas.

1. Evidencia relacionada con el contenido

La **validez de contenido** se refiere al grado en que un instrumento refleja un dominio específico de contenido de lo que se mide. Es el grado en el que la medición representa al concepto o variable medida (Bohrnstedt, 1976). Por ejemplo, una prueba de operaciones aritméticas no tendrá validez de contenido si incluye sólo problemas de resta y excluye problemas de suma, multiplicación o división. 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.

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

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.

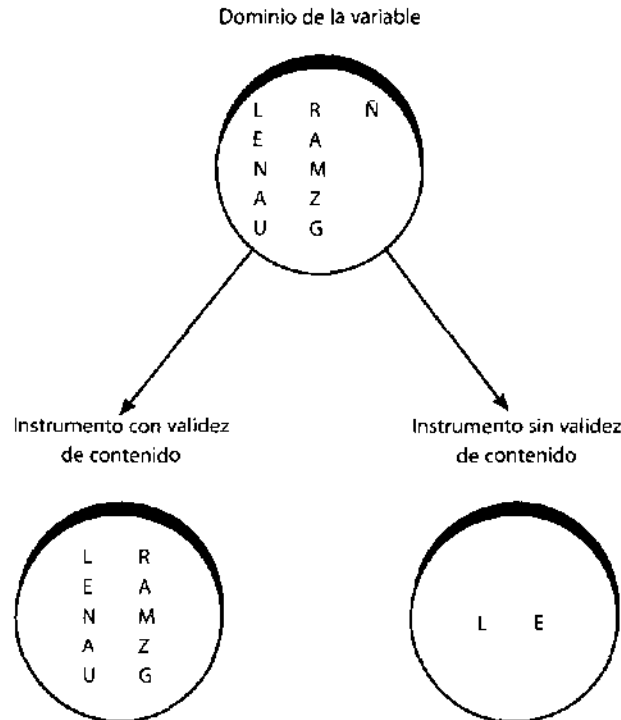


Figura 9.2 Ejemplo de un instrumento de medición con validez de contenido frente a otro que carece de ella.

El dominio de contenido de una variable normalmente es 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 constituido 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 2004. 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: moral, apoyo de la dirección, innovación, percepción de la empresa—identidad-identificación, comunicación, percepción del desempeño, motivación intrínseca, autonomía, satisfacción general, liderazgo, visión y 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. Esto se muestra en la figura 9.3.

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?*

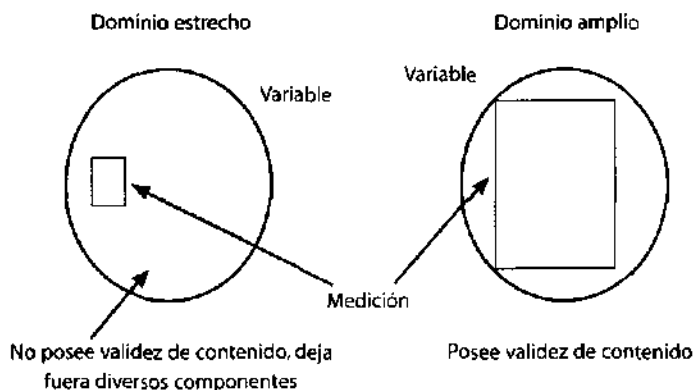


Figura 9.3 Ejemplo de dominio estrecho y amplio.

2. Evidencia relacionada con el criterio

La **validez de criterio** establece la validez de un instrumento de medición al compararla con 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

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

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, 2005). 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 un instrumento para medir el sentido de vida de acuerdo con la visión de V. Frankl. Para aportar evidencia de validez en relación con su instrumento —denominado Test Celaya— 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 Elizabeth 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.

Otro ejemplo fue el de Hernández Sampieri (2005), quien utilizó un criterio conductual para actualizar la validación de una escala de motivación intrínseca: el Cuestionario de reacción a tareas, en su versión en español. Diversos autores como Deci, Koestner y Ryan (1999) sostienen que el nivel de motivación intrínseca hacia una tarea está relacionado positivamente con el grado de persistencia adicional en el desarrollo de la tarea (por ejemplo, los empleados con mayor motivación intrínseca son los que suelen quedarse por más tiempo, una vez que concluye su jornada). En consecuencia, la predicción teórica es que a mayor motivación intrínseca habrá mayor persistencia adicional en la tarea. Entonces, el mencionado autor administró el cuestionario de motivación intrínseca a un grupo de trabajadores, de este modo, también determinó su persistencia adicional en el trabajo. Ambas mediciones estuvieron correlacionadas de forma positiva y sustancial, entonces se aportó evidencia para la validez de criterio del Cuestionario de reacción a tareas, versión en español (a la validez para medir la motivación intrínseca).

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 comparando 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 (Bostwick y Kyte, 2005). 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, 1989).

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 (Bostwick y Kyte, 2005). 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.

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

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 del marco teórico).
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. Esto se representa en la figura 9.4.

Veamos la validez de constructo con el ejemplo ya comentado de Hernández Sampieri (2005) sobre el clima organizacional.

EJEMPLO

El investigador aplicó un instrumento para evaluar al clima organizacional, el cual, recordemos que midió las siguientes variables: moral, apoyo de la dirección, innovación, percepción de la empresa-identidad-identificación, comunicación, percepción del desempeño, motivación intrínseca, autonomía, satisfacción general, liderazgo, visión y recompensas o retribu-

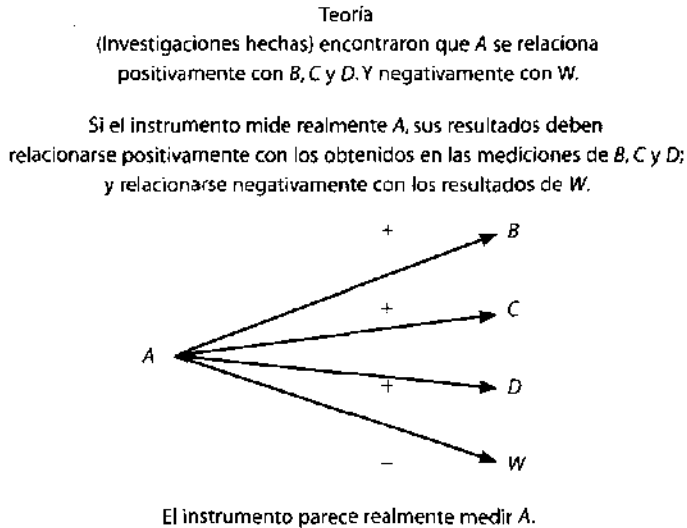


Figura 9.4 Presentación gráfica de un instrumento con validez de constructo.

ción. 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.

La teoría indica que tales dimensiones están fuertemente vinculadas y que se unen o funden entre sí para formar un constructo multidimensional denominado clima organizacional, y que además están relacionadas con el involucramiento en el trabajo y el compromiso organizacional.¹

Entonces, para aportar validez de constructo, las mediciones de las dimensiones deben estar asociadas entre sí y fundirse en un concepto más general (factor), asimismo, correlacionarse con mediciones del involucramiento y el compromiso. Para tal efecto se llevó a cabo un análisis de factores y otras pruebas multivariadas.² Éstas determinan la contribución de cada dimensión o concepto en todo el instrumento (cuando incluyen más de un componente y una variable). 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?*

¹ James y James, 1989; James *et al.*, 1990; James y James, 1992; James y McIntyre, 1996; Parker *et al.*, 2003; Rizzo, 2004...).

² No se comentan aquí tales pruebas, sino en el capítulo Análisis estadístico-multivariado de los datos del CD anexo, con el fin de no complicar el ejemplo al lector. Inclusive en tal capítulo se presentan los resultados de este caso.

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 (Bostwick y Kyte, 2005). 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, se comienza a comentar en torno a la *validez consecuente*, que se refiere a las secuelas sociales del uso e interpretación de una prueba (Mertens, 2005; Messick, 1995).

La validez total

La validez de un instrumento de medición se evalúa sobre la base de todos los tipos de evidencia. Cuanto 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.

Diferentes instancias académicas (entre las que podemos mencionar a la American Psychological Association, la American Educational Research Association, la Red de Profesores Investigadores Latinoamericanos Certificados), así como la mayoría de los autores contemporáneos, contemplan a la validez como un concepto unitario en el que se considera toda la evidencia acumulada con respecto a la interpretación de las puntuaciones de un instrumento para el propósito establecido (Mertens, 2005; Bostwick y Kyte, 2005). En este sentido, la validez no puede descansar en un solo tipo de evidencia (Messick, 1995, 1989 y 1988).

Validez total = validez de contenido + validez de criterio + 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 no mida lo que pretende). Por ello es requisito que el instrumento de medición demuestre ser *confiable y válido*. De no ser así, los resultados de la investigación no deben tomarse en serio.

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

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:

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

Sus resultados podrían visualizarse como en la figura 9.5, en la cual se vinculan la confiabilidad y la validez.

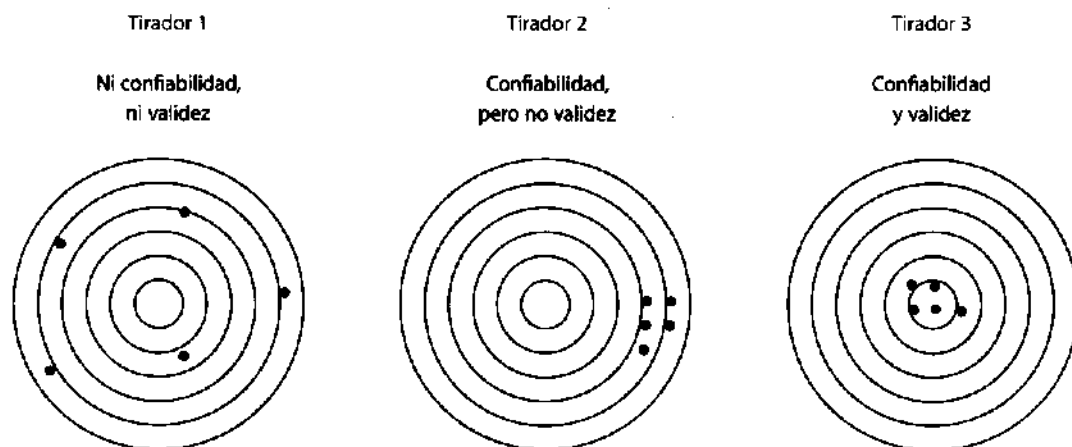


Figura 9.5 Representación de la confiabilidad y la validez.

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,³ 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, lo cual no debe existir en la investigación.

También a las y los investigadores experimentados les toma cierto tiempo desarrollar un instrumento de medición. Los construyen con cuidado y frecuentemente los desarrollan para que, cuando los necesiten con premura, se encuentren preparados para aplicarlos; pero no los improvisan. 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. Por ejemplo, seleccionar o generar un instrumento que mida la inteligencia, la personalidad, o los usos y las gratificaciones de la televisión para el niño, requiere amplios conocimientos en la materia, estar actualizados al respecto y revisar con cuidado la literatura correspondiente.

³ 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 "Recolección de los datos cuantitativos", segunda parte.

El segundo factor es que a veces se utilizan instrumentos desarrollados en el extranjero que no han sido validados para 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.

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 población 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 sujetos participantes y no éstos a él o ella, ya que es necesario brindarles todo tipo de facilidades.

Si este es el caso, sugerimos que se consulte a Mertens y McLaughlen (2004), en cuyo libro tienen un capítulo dedicado a la recolección de información con personas con capacidades diferentes o de culturas especiales. Asimismo, es recomendable revisar la página web de alguna asociación internacional como la American Psychological Association.

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.

El cuarto factor agrupa diversas cuestiones vinculadas con los estilos personales de los sujetos participantes (Bostwick y Kyte, 2005) tales como: deseabilidad 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 en domingo a las 7:00 AM o después de las 11:00 PM 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.

Otro 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 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 (Unrau, Grinnell y Williams, 2005). Al tratarse de cuestiones físicas las percepciones suelen compartirse (por ejemplo, la mayoría de las personas estarían de acuerdo que el agua de mar contiene sal o los rayos del sol queman), pero en tópicos 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.

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.

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 reduzcan 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.

... ¿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:

$$X = t + 0$$

$$X = t$$

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

Existen diversos procedimientos para calcular la confiabilidad de un instrumento de medición. Todos utilizan fórmulas que producen coeficientes de confiabilidad. La mayoría de estos coeficientes pueden oscilar entre cero y uno, donde un coeficiente de cero significa nula confiabilidad y uno representa un máximo de confiabilidad (confiabilidad total). Cuanto más se acerque el coeficiente a cero (0), mayor error habrá en la medición. Esto se ilustra en la tabla 9.2.

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

Tabla 9.2 Interpretación de un coeficiente de confiabilidad

Muy baja	Baja	Regular	Aceptable	Elevada
←				→
0				1
0% de confiabilidad en la medición (la medición está contaminada de error)				100% de confiabilidad en la medición (no hay error)

1. *Medida de estabilidad* (confiabilidad por *test-retest*). En este procedimiento un mismo instrumento de medición se aplica dos o más veces a un mismo grupo de personas, después de cierto periodo. Si la correlación entre los resultados de las diferentes aplicaciones es altamente positiva, el instrumento se considera confiable. Se trata de una especie de diseño panel. Desde luego, el periodo entre las mediciones es un factor a considerar. Si el periodo es largo y la variable susceptible de cambios, ello suele confundir la interpretación del coeficiente de confiabilidad obtenido por este procedimiento. Y si el periodo es corto las personas pueden recordar cómo respondieron en la primera aplicación del instrumento, para aparecer como más consistentes de lo que en realidad son (Bohrstedt, 1976). Algo así como la “administración de pruebas” en los experimentos.
2. *Método de formas alternativas o paralelas*. En este procedimiento no se administra el mismo instrumento de medición, sino dos o más versiones equivalentes de éste. Las versiones son similares en contenido, instrucciones, duración y otras características. Las versiones (casi siempre dos) se administran a un mismo grupo de personas simultáneamente o dentro de un periodo relativamente corto. El instrumento es confiable si la correlación entre los resultados de ambas administraciones es positiva de manera significativa. Los patrones de respuesta deben variar poco entre las aplicaciones. Una variación de este método es el de las formas alternas prueba-posprueba (Creswell, 2005), cuya diferencia reside en que el tiempo que transcurre entre la administración de las versiones es mucho más largo, que es el caso de algunos experimentos.
3. *Método de mitades partidas* (*split-halves*). Los procedimientos anteriores (medida de estabilidad y método de formas alternas) requieren cuando menos dos administraciones de la medición en el mismo grupo de individuos. En cambio, el **método de mitades partidas** necesita sólo una aplicación de la medición. Específicamente el conjunto total de ítems⁴ o reactivos se divide en dos mitades equivalentes y se comparan las puntuaciones o los resultados de ambas. Si el instrumento es confiable, las puntuaciones de las dos mitades deben estar muy correlacionadas. Un individuo con baja puntuación en una mitad tenderá a mostrar también una baja puntuación en la otra mitad. El procedimiento se diagrama en la figura 9.6.
4. *Medidas de consistencia interna*. Éstos son coeficientes que estiman la confiabilidad: a) *el alfa de Cronbach* (desarrollado por J.L. Cronbach) y b) *los coeficientes KR-20 y KR-21* de Kuder y

⁴ Un ítem es la unidad mínima que compone a una medición; es un reactivo que estimula una respuesta en un sujeto (por ejemplo, una pregunta, una frase, una lámina, una fotografía o un objeto de descripción).

Richardson (1937). El método de cálculo en ambos casos requiere una sola administración del instrumento de medición. Su ventaja reside en que no es necesario dividir en dos mitades a los ítems del instrumento, simplemente se aplica la medición y se calcula el coeficiente. La manera de calcular estos coeficientes y las correlaciones para todos los métodos aquí expuestos se explica en el siguiente capítulo sobre análisis de los datos cuantitativos y en el CD anexo.

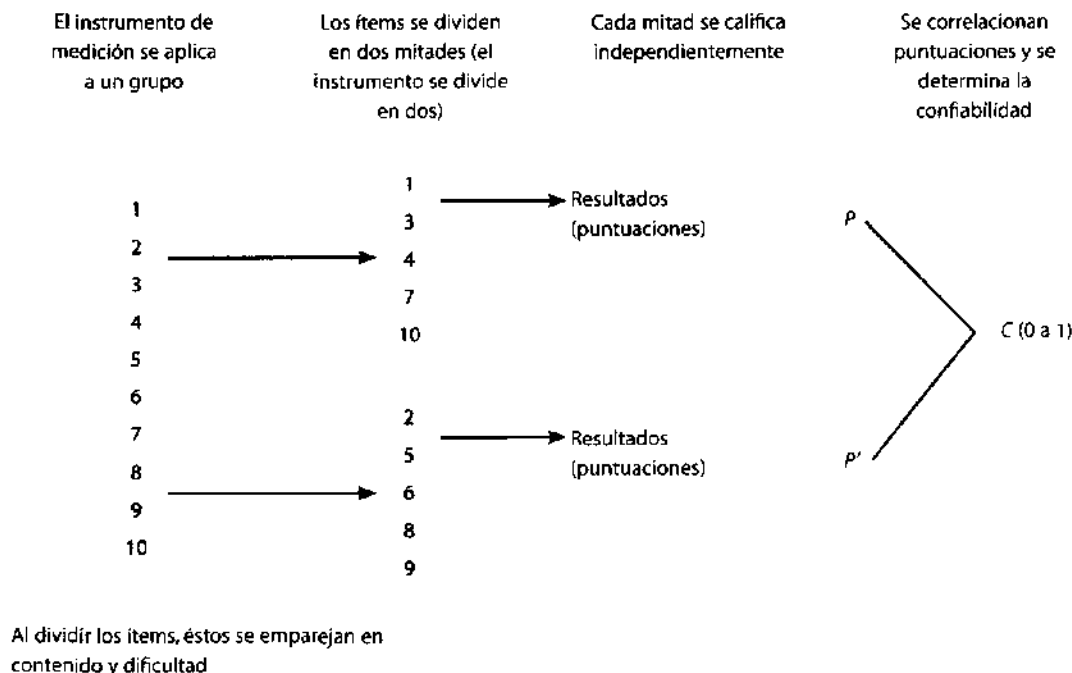


Figura 9.6 Esquema del procedimiento de mitades-partidas.

Con la finalidad de comprender mejor los métodos para determinar la confiabilidad vea la tabla 9.3.

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

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

Cálculo de la validez

Con respecto a la validez de contenido, que resulta más compleja que las otras (validez de criterio y de constructo) para calcularla: Primero, es necesario revisar cómo ha sido medida la varia-

Tabla 9.3 Aspectos básicos de los métodos para determinar la confiabilidad

Método	Número de veces que el instrumento es administrado	Número de versiones diferentes del instrumento	Número de participantes que proveen los datos	Inquietud o pregunta que contesta
Estabilidad (<i>test—retest</i>)	Dos veces en tiempos distintos.	Una versión.	Cada participante responde al instrumento dos veces.	¿Responden los individuos de una manera similar a un instrumento si se les administra dos veces?
Formas alternas	Dos veces al mismo tiempo o con una diferencia de tiempo muy corta.	Dos versiones diferentes, pero equivalentes.	Cada participante responde a cada versión del instrumento.	Cuando dos versiones de un instrumento son similares, ¿hay convergencia o divergencia en las respuestas a ambas versiones?
Formas alternas y prueba-pos-prueba	Dos veces en tiempos distintos.	Dos versiones diferentes, pero equivalentes.	Cada participante responde a cada versión del instrumento.	Cuando dos versiones de un instrumento son similares, ¿hay convergencia o divergencia en las respuestas a ambas versiones?
Mitades partidas	Una vez	Una fragmentada en dos partes equivalentes.	Cada participante responde a la única versión.	¿Son las puntuaciones de una mitad del cuestionario similares a las obtenidas en la otra mitad?
Medidas de consistencia interna (alfa y KR-20 y 21)	Una vez	Una versión	Cada participante responde a la única versión.	¿Las respuestas a los ítems del instrumento son coherentes?

ble 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) (Bohrnstedt, 1976), 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 (pero no de manera estadística sino conceptual).

La *validez de criterio* se estima al correlacionar la medición con el criterio externo (puntuaciones de su 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 9.7.⁵

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 capítulo Análisis estadístico-multivariado de los datos del CD anexo.



Figura 9.7 Ejemplo para el cálculo de la validez de criterio.

... ¿Qué procedimiento se sigue para construir un instrumento de medición?

Existen diversos tipos de instrumentos de medición, cada uno con características diferentes. Sin embargo, el procedimiento general para construirlos es semejante. Éste se resume en el diagrama de la figura 9.8 y corresponde a la parte del plan de recolección que responde a la pregunta: ¿a través de qué método vamos a recolectar los datos?

Revisemos ahora cada una de estas fases.

Fase 1: Redefiniciones fundamentales

A estas alturas del proceso de investigación cuantitativa supuestamente debemos tener claridad sobre los conceptos o variables de la investigación (se definieron en el planteamiento, se volvieron a considerar durante o después de la revisión de la literatura, se evaluaron una vez más al formular o no formular las hipótesis y seguimos teniéndolas en mente durante todo el proceso). Sin embargo, al comenzar a desarrollar el o los instrumentos de medición, vale la pena una última

⁵ Las pruebas de correlación se presentan en el siguiente capítulo.

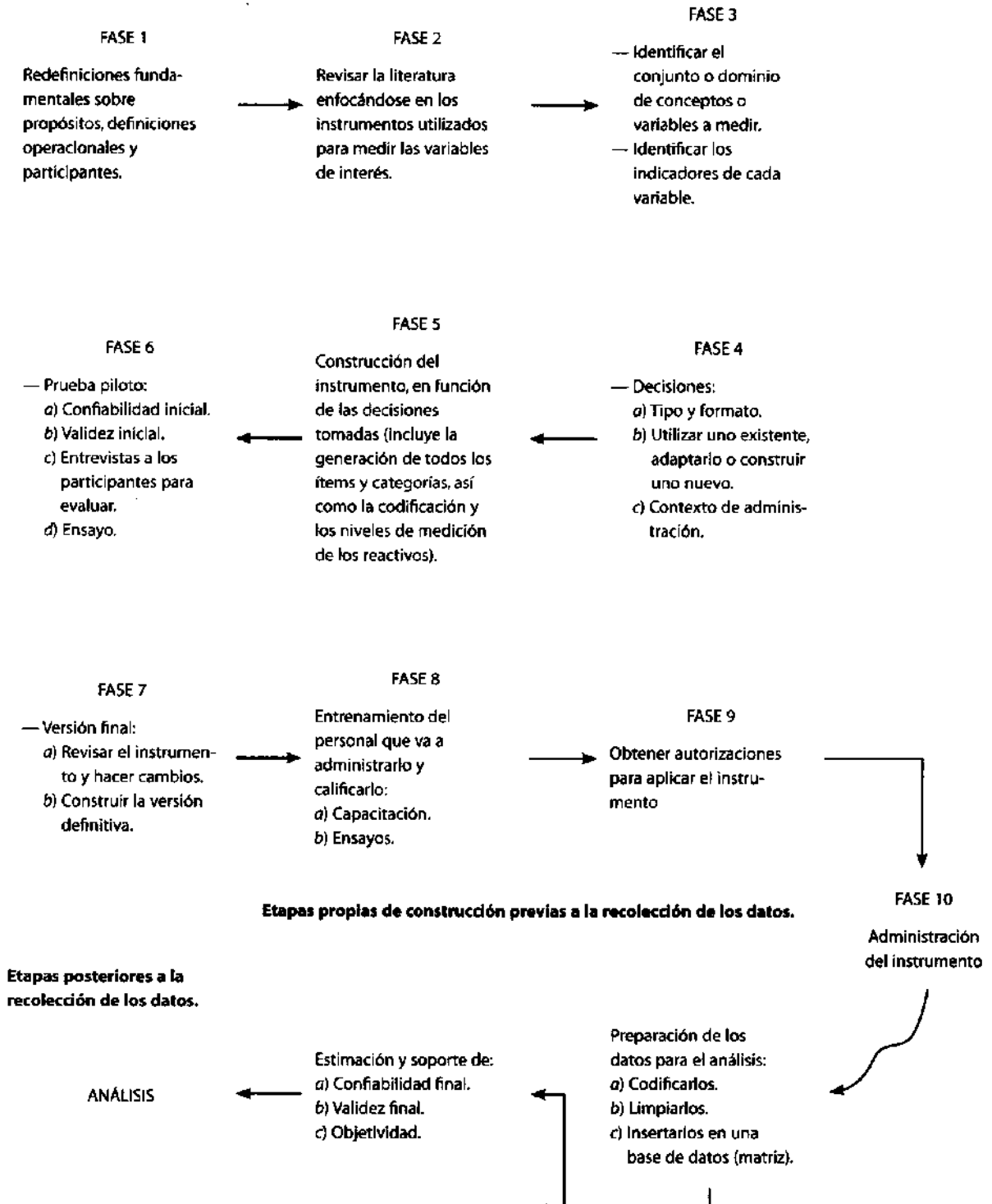


Figura 9.8 Proceso para construir un instrumento de medición.

reflexión para cerciorarnos de cuáles son nuestras variables y algunas otras precisiones: Para tal efecto, sugerimos contestar las siguientes preguntas:

¿Qué va a ser medido (identificación y listado de variables)?

Variable o concepto 1: _____

Variable o concepto 2: _____

Variable o concepto 3: _____

Variable o concepto k : _____

¿Qué o quiénes van a ser medidos? _____

¿Cuándo (fechas precisas)? _____

¿Dónde? (Lugar específico, por ejemplo, en el caso mencionado de los trabajadores de laboratorios farmacéuticos de León, qué empresas en concreto). _____

¿Nuestro propósito (coherente con el planteamiento) al recolectar los datos es? _____

¿Nuestras definiciones operacionales son?

Variable o concepto 1: _____

Variable o concepto 2: _____

Variable o concepto 3: _____

Variable o concepto k : _____

¿Qué tipo de datos queremos obtener? (Por ejemplo: respuestas verbales, conductas observadas, respuestas escritas, estadísticas, proyecciones internas, marcas de un sistema electrónico, etcétera). _____

Como resultado de responder a las preguntas, se tienen las redefiniciones más importantes para proseguir.

Fase 2: Revisión enfocada en la literatura

Si al elaborar el marco teórico efectuamos una revisión adecuada de la literatura, esta fase suele ser innecesaria. Pero, puede ocurrirle a quien se inicia en la investigación, que al consultar las fuentes primarias, no se haya enfocado en los instrumentos de medición utilizados. En esta situación, resulta indispensable volver a revisar la literatura, centrándose en tales instrumentos, lo cual resulta sencillo, porque —supuestamente— tenemos las referencias a la mano.

Al evaluar los diferentes instrumentos utilizados en estudios previos, conviene identificar —en cada caso— el propósito del instrumento tal como fue concebido por el investigador o investigadores, y analizar si su propósito es similar al nuestro. Éste es el primer punto que tomamos en cuenta para ver qué instrumentos nos pueden ser útiles. Asimismo, se recomienda considerar los siguientes elementos:

- Las variables que mide el instrumento o herramienta: ¿cuántas y cuáles?
- Los dominios de contenido de las variables (dimensiones, componentes o indicadores): ¿cuáles? (Mertens, 2005)
- La muestra: ¿a quién o quiénes fue administrado?, ¿qué perfil o características poseen?
- Las condiciones de administración: ¿en qué lugar y contexto fue utilizado?, ¿cómo y cuándo fue administrado?, ¿por quién o quiénes se aplicó?, ¿eran individuos con capacidades y entrenamiento para implementarlo, calificarlo e interpretarlo, ¿qué tiempo requiere para ser administrado, codificado y calificado?
- La información sobre la confiabilidad: ¿de qué tipo?, ¿cómo fue establecida?, ¿qué procedimientos se utilizaron para calcularla?, ¿cómo se reportó?, ¿los resultados son aceptables y significativos?
- La información sobre la validez: ¿qué evidencia de contenido, criterio y constructo se aportó?, ¿se estableció suficiente evidencia sobre la validez total?
- La información sobre la objetividad: ¿se aplicó de manera estandarizada?
- Los recursos necesarios para administrarlo: ¿qué recursos se requieren para aplicarlo? (por ejemplo: informáticos, el costo).

Sobre la base de tales elementos, se seleccionan uno o varios instrumentos que pudieran ser apropiados para nuestra investigación.

Fase 3: Identificación del dominio de las variables a medir y sus indicadores

Esta fase es prácticamente simultánea a la anterior y simplemente se refiere a la definición del dominio de las variables que habremos de medir, fundamentándonos en la revisión de otros instrumentos. Es decir, señalar con precisión los componentes, dimensiones o factores que teóricamente integran a la variable. Incluso, se sugiere hacerlo gráficamente (ver figura 9.9).

Asimismo, establecer los indicadores de cada dimensión. En el caso de la “satisfacción en torno a las condiciones de trabajo”, éstos podrían ser: jornada, horarios de entrada, salida y alimentación, lugar, transporte, iluminación, servicio médico y vacaciones. Al tratarse de la “satisfacción sobre el propio desempeño en el trabajo”: productividad, eficiencia, eficacia, calidad en el servicio, desperdicio y calidad factual.

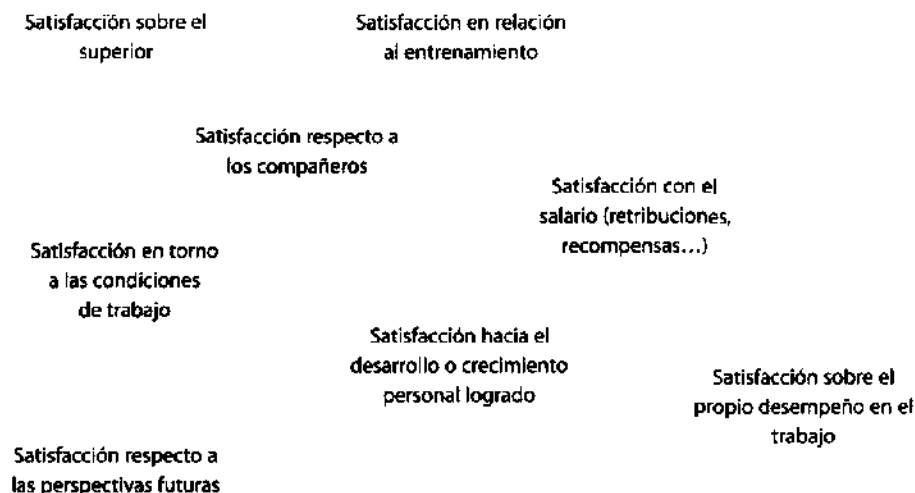


Figura 9.9 Ejemplo con la variable satisfacción laboral e indicadores.

Fase 4: Decisiones clave

Una vez evaluados varios instrumentos (al menos un par) utilizados previamente, es necesario tomar una serie de decisiones fundamentales, que agrupamos en tres rubros:

- 1) Utilizar un instrumento de medición ya elaborado, adaptarlo o desarrollar uno nuevo.

En el primer caso, debemos elegir un instrumento que obviamente se encuentre disponible ("esté a la mano" y podamos utilizarlo sin violar ningún derecho de autoría) y sea posible su aplicación (al respecto, debemos cuestionarnos si en realidad podemos aplicarlo). Asimismo debemos seleccionar el que *a)* se haya generado lo más recientemente posible; *b)* se cite con mayor frecuencia y amplitud en la literatura; *c)* reciba mejores evaluaciones por expertos de asociaciones académicas y profesionales; *d)* demuestre mayor confiabilidad, validez y objetividad; *e)* se pueda aplicar más a nuestro planteamiento y propósito de recolección de los datos; *f)* se adecue mejor a nuestra muestra y contexto, y *g)* empate en mayor medida con las aptitudes que tenemos (que estemos más calificados para utilizarlo). Solamente deben seleccionarse instrumentos que reporten confiabilidad, validez y objetividad; ya que no es posible confiar en una forma de medir que carezca de este tipo de evidencia clara y precisa. Cabe señalar que algunos instrumentos se han validado para contextos latinoamericanos y españoles. Por ejemplo, la ya comentada Escala Clima-UNI con 96 ítems para medir el clima organizacional, puede ser aplicada tal cual al medio laboral mexicano (Hernández Sampieri, 2005). Su confiabilidad promedio es de 0.95 y posee evidencia de validez de todo tipo. Desde luego, si se quisiera administrar en otra nación latinoamericana, debe volver a validarse; y en el caso de España, tendría que adecuarse. Si se trata de una herramienta escrita (como un cuestionario), es ineludible revisar la redacción.

En la segunda situación, el instrumento elegido se adapta al ambiente de nuestro estudio. Es decir, se realizan algunos ajustes en la presentación, instrucciones, ítems (preguntas, categorías

de observación, etc.). También se revisa cuidadosamente el lenguaje y la redacción. Éste es el caso de instrumentos que se traducen y validan. Por ejemplo, el instrumento Children's Knowledge of Abuse Questionnaire-Revised (CKAQ-R), versión en español, fue adaptado de su homólogo, el cual fue generado en Estados Unidos y se tradujo como Cuestionario del Conocimiento del Abuso Infantil-Revisado.

Si ninguna de las dos opciones anteriores resultan satisfactorias, entonces tenemos que desarrollar uno propio. Éste puede elaborarse a partir de varios instrumentos previos, de nuevas concepciones e ítems, o bien, una mezcla de éstas. Para tal efecto, idealmente debe generarse un conjunto de ítems o categorías por los indicadores de cada dimensión, con el apoyo de un grupo de expertos (en una o varias sesiones del tipo denominado "tormenta de ideas". Posteriormente, someter a los ítems a evaluación de otros expertos y redactar la introducción e instrucciones. Por ejemplo, Hernández Sampieri (2005), construyó un instrumento mediante la selección de ítems de varios cuestionarios y escalas desarrolladas en Estados Unidos, y otras más que han sido aplicadas en México, además de incluir nuevos ítems.

En la tabla 9.4 se presentan ejemplos de cómo se transita de la variable a la dimensión, de ésta al indicador, hasta llegar a los ítems y sus categorías.

Tabla 9.4 Ejemplo de desarrollo de ítems⁶

Estudio de las preferencias de los jóvenes para divertirse (ejemplo sencillo)⁷

Variable	Dimensión	Indicador	Ítems
Preferencia 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 agrade). 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) Otro (especificar).
	Actividad nocturna en fin de semana.	Jerarquía de preferencias de actividades en viernes y sábado.	Mismas categorías u opciones de respuesta.

(continúa)

Tabla 9.4 Ejemplo de desarrollo de ítems (continuación)

Estudio de las preferencias de los jóvenes para divertirse (ejemplo sencillo)			
Variable	Dimensión	Indicador	Ítems
	Actividad nocturna en domingo.	Jerarquía de preferencia de actividades en domingo.	Mismas categorías u opciones de respuesta.
Investigación sobre clima organizacional			
Variable	Dimensión	Indicador	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.	- 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. - 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

(continúa)

Tabla 9.4 Ejemplo de desarrollo de ítems (*continuación*)

Investigación sobre clima organizacional			
Variable	Dimensión	Indicador	Ítems
	Autonomía	Grado de libertad percibida para tomar decisiones y realizar el trabajo.	– 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. – 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.	– 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).
Investigación sobre la relación vendedor-comprador industrial			
Variable	Dimensión	Indicador	Ítems
Percepción del comprador sobre la coordinación entre su organización y las empresas proveedoras	Coordinación de conflictos	Esfuerzo de ambas partes por no provocar conflictos.	¿Cuánto se esfuerza su empresa para no provocar problemas con sus proveedores? 1) Se esfuerza al mínimo posible. 2) Se esfuerza poco. 3) Se esfuerza medianamente. 4) Se esfuerza mucho. 5) Se esfuerza al máximo posible.

(continúa)

Tabla 9.4 Ejemplo de desarrollo de ítems (*continuación*)

Investigación sobre la relación vendedor-comprador industrial

Variable	Dimensión	Indicador	Ítems
			¿Cuánto se esfuerzan sus proveedores para no provocar problemas con su empresa? (Mismas opciones de respuesta que el ítem anterior).
	Coordinación de no interferencia	Esfuerzo de ambas partes por no interferir en la otra empresa.	¿Cuánto se esfuerza su empresa para no interferir en los deberes y las responsabilidades de sus proveedores? 5) Se esfuerza al máximo posible. 4) Se esfuerza mucho. 3) Se esfuerza medianamente. 2) Se esfuerza poco. 1) Se esfuerza al mínimo posible.
			¿Cuánto se esfuerzan sus proveedores para no interferir con los deberes y las responsabilidades de su empresa? (Mismas opciones de respuesta que el ítem anterior).

⁶ En los ejemplos de la tabla sólo se incluyen unos cuantos ítems o reactivos por cuestiones de espacio. Son ejemplos muy resumidos.

⁷ Nótese que fin de semana pudiera definirse de varios modos: de jueves a domingo, viernes a domingo, viernes y sábado. Asimismo, si las opciones de respuesta no reflejan la realidad del joven, el instrumento no está adecuado al contexto y entonces estamos en problemas de operacionalización de la variable. Los expertos aquí serían jóvenes del país en el que lo vamos a aplicar.

2) Resolver el tipo de instrumento y formato.

La segunda decisión es simultánea a la primera cuando se utiliza un instrumento ya elaborado o adaptado, pues éste ya posee sus características y tiene un formato propio. Pero cuando se desarrolla uno nuevo, el investigador debe decidir de qué tipo será (cuestionario, hoja de observación, escala de actitudes, prueba estandarizada, etc.) y cuál será su formato, esto depende de sus características físicas (tamaño físico, colores, tipo de fuente, especificaciones de papel y aspectos similares). Por ejemplo, no es lo mismo un cuestionario escrito en papel que un cuestionario en un disco compacto (CD) o en una página web.

3) Determinar el contexto de administración.

Esto implica decidir en qué espacio físico se implementa (si en un aula de capacitación, en un salón de clases, en la calle, en los hogares, de manera telefónica, en una cámara de Gesell, etc.). Desde luego, la naturaleza de los datos buscados y el tipo de instrumento elegido nos ayudan en ello, ya sabemos que si deseamos datos escritos y vamos a utilizar un cuestionario, éste puede administrarse en distintos contextos. Mientras que si nuestro interés es observar, lo podemos hacer en un lugar público, en un salón o en una cámara de Gesell, por ejemplo. Cuando se revisen diferentes instrumentos, se profundizará en el tema.

Fase 5: Construcción del instrumento

Esta etapa implica la generación de todos los ítems o reactivos y categorías del instrumento, como se presentó de manera resumida en la tabla 9.4. Esta construcción depende del tipo de instrumento y formato. Recordemos que un ítem es, por ejemplo, una pregunta de un cuestionario, una categoría de un sistema de observación o una estadística en un sistema de registro. Al final de esta etapa el producto es una primera versión del instrumento. Una parte muy importante de esta versión, es cómo vamos a registrar los datos y el valor que les habremos de otorgar, a esto se le conoce como “codificación”.

Generación de ítems Debemos asegurarnos de tener un número suficiente de ítems para medir todas las variables en todas sus dimensiones.

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

Categoría	Codificación (valor asignado)
— Masculino	1
— Femenino	2

Así, Becky Costa en la variable género sería 2, Luis Gerardo Vera y José Ramón Calderón serían 1, Verónica Larios 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 9.5.

Tabla 9.5 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 9.4, 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 = otro.

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) 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

¿Tiene usted novia?

☐ 1 Sí

☐ 0 No

Pregunta no precodificada

¿Tiene usted novia?

☐ Sí

☐ No

Se incluye la codificación
(en este caso, números)

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 debido a que 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.

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.

Sexo { Masculino
Femenino

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

3 = Protestante

4 = Musulmán

5 = Otros

1 + 2 = 3

¿Un católico + un judío = un protestante?

(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 éstas 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:⁸

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 escala⁹ 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:

Presidente	10
Vicepresidente	9
Director general	8
Gerente de área	7
Subgerente o superintendente	6
Jefe	5
Empleado A	4

⁸ Duncan (1977).

⁹ Nam *et al.* (1965).

Empleado B	3
Empleado C	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 $10 \text{ (presidente)} \div 5 \text{ (jefe)} = 2 \text{ (empleado C)}$. 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.

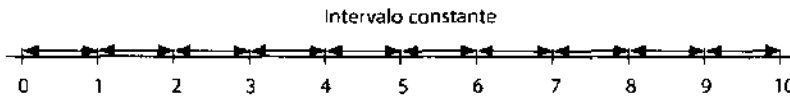


Figura 9.10 Escala con intervalos iguales entre categorías.

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 (ver figura 9.11).

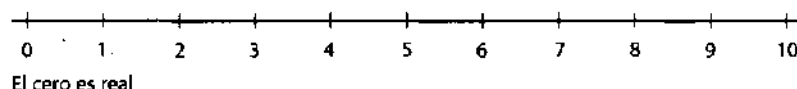


Figura 9.11 Ejemplo de escala para el nivel de medición de razón.

Ejemplos de estas mediciones serían la exposición a la televisión (en minutos), el número de hijos, la productividad, las ventas de un producto, el ingreso, los metros cuadrados de construcción, peso de materiales en kilogramos, distancia de alcance de una microonda, número de años en el sistema educativo, presión arterial, entre otros.

Desde luego, hay variables que pueden medirse en más de un nivel, según el propósito de medición. Por ejemplo, la variable “antigüedad en la empresa”:

<i>Nivel de medición</i>	<i>Categorías</i>
— De razón	En días (0 a K días)
— Ordinal	Bastante antigüedad Antigüedad regular Poca antigüedad

Ya desarrollamos una versión preliminar del instrumento, que habremos de probar.

Fase 6: Prueba piloto

Esta prueba consiste en administrar el instrumento a personas con características semejantes a las de la muestra objetivo de la investigación. Se somete a prueba no sólo el instrumento de medición, sino también las condiciones de aplicación y los procedimientos involucrados. Se analiza si las instrucciones se comprenden y si los ítems funcionan de manera adecuada, se evalúa el lenguaje y la redacción. Los resultados se utilizan para calcular la confiabilidad inicial y, de ser posible, la validez del instrumento de medición. La **prueba piloto** se realiza con una pequeña muestra (inferior a la muestra definitiva). Los autores aconsejamos que cuando la muestra sea de 300 o más se lleve a cabo la prueba piloto con entre 30 y 60 personas, salvo que la investigación exija un número mayor.

Por ejemplo, Hernández Sampieri (2005) sometió el cuestionario del clima organizacional a una prueba piloto en dos muestras de 30 y 32 participantes respectivamente. La confiabilidad inicial de estas pruebas fue de 0.91 y 0.92.

En ocasiones, el instrumento se somete a varias pruebas y se va depurando paulatinamente. Núñez (2001), generó una primera versión de su escala del sentido de vida (conformada por 135 reactivos), la aplicó a un grupo y realizó ajustes sobre los resultados; posteriormente, una segunda versión (180 ítems) fue administrada a otro grupo piloto similar al primero; se volvieron a efectuar cambios en el instrumento, hasta que llegó a una versión final, la cual contenía 99 reactivos.

Prueba piloto Consiste en administrar el instrumento a una pequeña muestra, cuyos resultados se usan para calcular la confiabilidad inicial y, de ser posible, la validez del instrumento.

Cabe señalar que cuando se desarrolla un nuevo instrumento, es conveniente incluir un elevado número de ítems, para elegir a los que más contribuyan a la confiabilidad, validez y objetividad de éste. Desde luego, sin excedernos, sin caer en redundancias excesivas, el número o cantidad de reactivos debe ser manejable. La inclusión de cada ítem debe estar justificada (Mertens, 2005). Por ello, se recomienda que la generación de reactivos se realice mediante una o varias sesiones con expertos (Gall, Gall y Borg, 2003), los cuales pueden ser profesores universitarios, consultores de empresas y especialistas que uno conozca. Cuando no encontramos expertos, entonces nuestra revisión de la literatura tiene que ser exhaustiva y profunda. Recalcamos que para ello es conveniente consultar el CD anexo en este tema.

Como ya se comentó, la prueba piloto evalúa todo el procedimiento de medición, por ejemplo, en el experimento de Naves y Poplawsky (1984), se evaluó la conducta de aceptación-avoidancia de los participantes hacia el supuesto individuo con capacidad mental distinta (que recordemos era un actor) mediante observación. Cada interacción se filmó y se analizó. Durante la prueba piloto se descubrió que la iluminación era insuficiente y no se podían registrar ciertas conductas no verbales, entonces se mejoró.

Parte fundamental de la prueba piloto consiste en charlar con los participantes para recoger sus opiniones con respecto al instrumento y al contexto de aplicación (por ejemplo, lo consideran largo o corto, comprensible, obstrusivo o no, etc.). Una vez, al comenzar una prueba piloto de un instrumento para evaluar al superior inmediato, uno de los sujetos inquirió: "En el cuestionario se pregunta sobre nuestro jefe, pero, ¿a cuál de ellos se refieren? Porque en estos momentos en el área de producción tenemos dos superiores, el de siempre, nuestro supervisor regular, pero cómo estamos fabricando un nuevo modelo de automóvil, también tenemos un superior de este otro proyecto". Sin la prueba piloto, al levantar los datos no se hubiera sabido a qué superior se referían, la confusión hubiera sido inevitable.

A continuación incluimos algunas recomendaciones para tal prueba en el caso de instrumentos escritos (cuestionario, examen o prueba estandarizada), en la tabla 9.6.¹⁰

Tabla 9.6 Recomendaciones para la prueba piloto en instrumentos escritos

- Es aconsejable que se dispongan espacios para comentarios de los sujetos participantes.
- Solicite a los participantes que señalen ambigüedades, opciones o categorías no incluidas, términos complejos y redacción confusa.
- Tome en cuenta y evalúe lo escrito por los participantes durante todo el proceso.
- Esté atento a que no se presenten "patrones tendenciosos" por efecto del instrumento en: a) el orden de las preguntas, b) el orden de las opciones de respuesta, c) tendencias en las respuestas (por ejemplo: que siempre tiendan a estar de acuerdo o en contra de todo, sin que realmente sea su opinión; o que respondan "no sé" a una buena parte de las preguntas; que dejen de contestar a varios ítems).
- Simule las condiciones reales de administración de la mejor manera posible (por ejemplo, si el cuestionario va a enviarse por correo, la prueba piloto debe hacerse por este medio: se les manda y posteriormente, pero de manera inmediata, se les reúne en grupo para comentar; o al menos, se simulan las condiciones en el grupo: a los participantes se les entrega el cuestionario en un sobre postal, el cual abren y leen).

¹⁰ Algunas de estas sugerencias fueron adaptadas de Gall, Gall y Borg (2003) y Mertens (2005).

Fase 7: Versión final

Así, con los resultados de la prueba, el instrumento de medición preliminar se modifica, ajusta y mejora (se quitan o agregan ítems, se cambian palabras, se otorga más tiempo para responder, etc.). Se tiene la versión final para administrar, la cual incluye un diseño gráfico atractivo para los participantes.

Fase 8: Entrenamiento del personal

Las personas del equipo de investigación que van a administrar y/o calificar el instrumento, ya sea que se trate de entrevistadores, observadores, encuestadores o que dirijan las sesiones para aplicarlo, deben recibir un entrenamiento en varias cuestiones:

- a) La importancia de su participación y su papel en el estudio.
- b) El propósito del estudio.
- c) Las personas e instituciones que patrocinan y/o realizan la investigación.
- d) La calendarización y tiempos del estudio (fechas clave de entrega de información).
- e) La confidencialidad que deben guardar de la investigación y sobre todo de los participantes. Ésta no puede romperse bajo ninguna circunstancia.
- f) La manera cómo deben vestirse si van a entrevistar o tener contacto con los participantes.
- g) El respeto y sensibilidad que deben tener hacia los participantes (por ejemplo, no discriminar ni utilizar lenguaje que pudiera resultar ofensivo).
- h) No discutir ni molestarse con los y las participantes (ellos y ellas nos están haciendo un favor).
- i) Establecimiento de confianza.
- j) Uso de gafetes y cartas que los identifiquen.
- k) El instrumento de medición (estructura, orden, ítems, instrucciones para el personal y para los sujetos participantes) y las condiciones de administración.

Durante el entrenamiento, el personal que habrá de aplicar y/o calificar las respuestas o mediciones, debe realizar varios ensayos a la vista de todos (de entrevista, observación, aplicación de prueba, etc.), para así corregir aquellas situaciones que puedan afectar la objetividad y el correcto desempeño en la recolección de los datos.

El entrenamiento puede llevarse a cabo mediante una o más sesiones, según sea el caso. Se recomienda —al menos— una reunión previa al inicio de la recolección de los datos y una después de que se terminó el levantamiento de la información. Durante la sesión o sesiones se motiva al personal, se escuchan sus sugerencias y se responden inquietudes.

Fase 9: Autorizaciones

Antes de recolectar los datos, es necesario gestionar varias autorizaciones. En primer término, por parte de los representantes de las organizaciones a las cuales pertenecen los participantes (directivos y líderes sindicales en las empresas, directores y maestros en las escuelas, funcionarios gubernamentales, etc.). Para este fin conviene producir una breve presentación del estudio

en algún formato multimedia o de video (que pueda reproducirse utilizando un programa para PC, un DVD, u otros) y presentar el protocolo o proyecto escrito que de cualquier manera debemos elaborar. En el CD anexo, se muestran ejemplos de ambos elementos.

En segundo término, es conveniente obtener el consentimiento de los propios participantes, particularmente por un medio escrito. Desde luego, en encuestas en la vía pública o en los hogares, a veces esto resulta complicado y la autorización suele ser verbal. El asunto es más delicado cuando se trata de menores de edad y personas con capacidades diferentes. En este caso, el consentimiento es de los tutores y de las propias personas que se integrarán a la muestra. No obtenerlo o lograrlo mediante engaños es una clara *violación a la ética de la investigación*. No puede filmarse a alguien sin su autorización.

Fase 10: Administración del instrumento

Éste es el momento de la verdad, todo nuestro trabajo conceptual y de planeación se confronta con los hechos. La administración depende del tipo de instrumento que hayamos seleccionado y la experticia de todo el equipo que participa. Por ello, es momento de revisar las principales clases de instrumentos para medir. Debemos aclarar que una cuestión es el tipo de instrumento en sí y otra el contexto de administración. En la literatura sobre recolección de datos cuantitativos todavía hay cierta confusión al respecto. Por ejemplo, ciertos autores desarrollan el tema de los “cuestionarios” en las secciones de *surveys* o encuestas, otros en el apartado sobre entrevistas, algunos más los consideran en sí mismos como un tema o los incluyen en medición. Debido a esto, proponemos ver cada instrumento y dentro de éste, los contextos más comunes en que se usan.

Los instrumentos que serán revisados en este capítulo son: cuestionarios,¹¹ escalas de actitudes, registros del contenido (análisis de contenido) y observación. En el capítulo *Recolección de los datos cuantitativos, segunda parte*, que encontrará en el CD anexo, se comentarán los siguientes: 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.

¿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.

¹¹ Las entrevistas se plantean como un contexto en el cual pueden ser administrados los cuestionarios.

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. 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.

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 puede haber?

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*.

Preguntas cerradas

Las **preguntas cerradas** contienen categorías u opciones de respuesta que han sido previamente delimitadas. Es decir, se presentan a los participantes las posibilidades de respuesta, 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.

¿Estudia usted actualmente?

☐ Sí

☐ No

¿Durante la semana pasada vio la final de la Liga de Campeones de Europa?

☐ Sí

☐ No

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?

☒ Presidente/Director general

- Vicepresidente/Director corporativo
- Subdirector/Director/Gerente
- Subgerente/Superintendente
- Coordinador
- Jefe de área
- Supervisor
- Empleado
- Obrero
- Otro

Si usted tuviera elección, ¿preferiría que su salario fuera de acuerdo con su productividad en el trabajo?

- ☐ Definitivamente sí
- ☐ Probablemente sí
- ☐ No estoy seguro
- ☐ Probablemente no
- ☐ Definitivamente no

Como puede observarse, en las preguntas cerradas las categorías de respuesta son definidas *a priori* por el investigador y se le muestran al sujeto encuestado, quien debe elegir la opción que describa más adecuadamente su respuesta. Gambara (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 sujeto puede seleccionar más de una opción o categoría de respuesta.

EJEMPLO

Esta familia tiene:

- ☐ Radio
 - ☐ Televisión
 - ☐ Reproductor de DVD
 - ☐ Teléfono (línea telefónica en casa)
 - ☐ Computadora
 - ☐ Internet
 - ☐ Teléfono celular
-

Algunos sujetos marcarán una, dos, tres, cuatro o más opciones de respuesta. Las categorías no son mutuamente excluyentes. Otro ejemplo sería la siguiente pregunta:

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

De la sala de lectura:

- ☐ No entré
- ☐ A consultar algún libro
- ☐ A consultar algún periódico
- ☐ A consultar alguna revista
- ☐ A estudiar
- ☐ A utilizar la computadora
- ☐ A buscar a alguna persona
- ☐ Otros (especificar):

De la hemeroteca:

- ☐ No entré
- ☐ A consultar algún periódico
- ☐ A utilizar los reproductores de video (casete/DVD)
- ☐ A estudiar
- ☐ A hacer trabajos
- ☐ A sacar copias
- ☐ A leer algún libro
- ☐ A consultar vía Internet
- ☐ Otros (especificar):

Del mostrador de préstamos:

- ☐ No fui
- ☐ A solicitar algún libro
- ☐ A solicitar alguna tesis
- ☐ A solicitar algún periódico/revista
- ☐ A solicitar un CD/DVD
- ☐ A solicitar equipo audiovisual/de cómputo
- ☐ A solicitar asesoría para localización de material
- ☐ Otros (especificar):

En ocasiones, el encuestado tiene que jerarquizar opciones. Por ejemplo: ¿quién te atrae más de tus compañeras del salón de clases?, ¿cuál en segundo lugar?, ¿cuál en tercer lugar?

- ☐ Rocío
- ☐ Lucía
- ☐ Laura
- ☐ Bertha
- ☐ Mariana
- ☐ Otra

O bien, designar un puntaje a una o diversas cuestiones.

EJEMPLO

¿CUÁNTO LE INTERESA DESARROLLAR?

(Indique de 1 a 10 en cada caso según sus intereses. Máxima calificación = 10.)

- ___ Administración de sueldos y compensaciones
 - ___ Salud, seguridad e higiene
 - ___ Administración y negociación de contratos
 - ___ Relaciones con sindicatos
 - ___ Habilidades de comunicación ejecutiva
 - ___ Programas y procesos sobre calidad/productividad
 - ___ Calidad de vida en el trabajo
 - ___ Teoría de la organización
 - ___ Administración financiera
 - ___ Desarrollo organizacional/innovación
 - ___ Técnicas de investigación organizacional
 - ___ Estructura organizacional (tamaño, complejidad, formalización)
 - ___ Sistemas de información y control
 - ___ Auditoría administrativa
 - ___ Planeación estratégica
 - ___ Sistemas de computación
 - ___ Mercadotecnia y comercialización
 - ___ Otros (especificar):
-

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

¿Qué porcentaje de los alumnos de la Universidad de Navarra trabaja?

En algunas más, el sujeto 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.

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

☺ 100 – Completamente enamorada.

99

98

.

.

.

80

70

60

☹ 50

.

.

.

20

10

.

.

2

1

☹ 0 – Nada enamorada

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.

EJEMPLO

¿Por qué asiste a psicoterapia?

¿Qué opina del programa de televisión *Los cazadores*?

¿De qué manera la dirección de la empresa ha logrado la cooperación del sindicato para el proyecto de calidad?

Desde su punto de vista, ¿cómo definiría la cultura fiscal?

¿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. Éstos 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 es enviado 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 (Vinnesa, 2005).

La *principal desventaja de las preguntas cerradas* reside en que limitan las respuestas de la muestra y, en ocasiones, ninguna de las categorías describe con exactitud lo que las personas tienen en mente; no siempre se captura lo que pasa por la cabeza de los sujetos. Su redacción exige mayor laboriosidad y un profundo conocimiento del planteamiento por parte del investigador o investigadora (Vinnesa, 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 los cuales se les administrarán conocen y comprenden las categorías de respuesta. Por ejemplo, si preguntamos qué canal de televisión es el preferido, determinar las opciones de respuesta y que los sujetos las comprendan es muy sencillo. Pero si preguntamos sobre las razones y los motivos que provocan esa preferencia, señalar las opciones es algo bastante 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). 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 a 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.

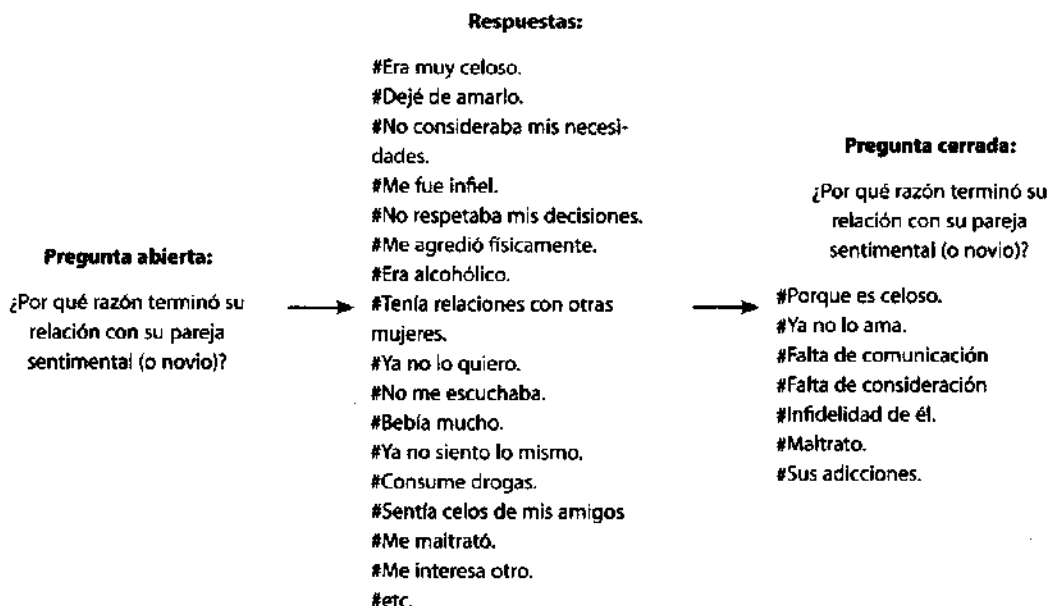


Figura 9.12 Ejemplo del paso de una pregunta abierta a la elaboración de una pregunta cerrada.

¿Una o varias preguntas para medir una variable?

En ocasiones sólo basta una pregunta para recolectar la información necesaria sobre la variable a medir. Por ejemplo, para medir el nivel de escolaridad de una muestra basta con preguntar: ¿Hasta qué año escolar cursó? o ¿cuál es su grado máximo de estudios? En otras ocasiones se requiere elaborar varias preguntas para verificar la consistencia de las respuestas. Por ejemplo, el nivel de ingresos familiar se mide cuestionando a los miembros de la familia que trabajan: ¿aproximadamente cuál es su nivel mensual de ingresos? y preguntando a la “cabeza o jefe de ésta”: ¿cuántos focos eléctricos tiene aproximadamente en su casa?¹² Además de preguntar sobre propiedades e inversiones, etcétera.

Al respecto, es recomendable *hacer solamente las preguntas necesarias* para obtener la información deseada o medir la variable. Si una pregunta es suficiente, no es necesario incluir más. No tiene sentido. Si se justifica hacer varias preguntas, entonces es conveniente plantearlas en el cuestionario. Esto último ocurre con frecuencia en el caso de variables con varias dimensiones o componentes a medir, en los cuales se incluyen varias preguntas para medir las distintas dimensiones. Se tienen varios indicadores.

EJEMPLO

La empresa Comunicometría, S.C., realizó una investigación para la Fundación Mexicana para la Calidad Total, A.C. (1988), con el propósito de conocer 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.

¹² 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 (Comunicometría, 1988).

- 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 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).¹³

Preguntas o ítems:	Completamente (mucho)	Completamente	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

¹³ El ejemplo contiene solamente algunas de las preguntas de la escala original.

¿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 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).

EJEMPLOS

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 no puede darse la precodificación, 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 sujeto encuestado: género, edad, nivel socioeconómico, estado civil, escolaridad (nivel de estudios), 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? y ¿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 (Corbetta, 2003), 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 significa 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 (2002) 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). Recuérdese que es ineludible tomar en cuenta su nivel educativo y el socioeconómico, las palabras que maneja, etcétera.
- d) No pueden incomodar al sujeto encuestado, ni ser percibidas como amenazantes y nunca éste debe sentir que se le enjuicia. Debemos inquirir de manera sutil. Preguntas tales 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 sujeto 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?, expre-

sa 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 improductivos	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, con 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 el 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? 25% 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 9.7 sobre problemas al generar preguntas, adaptado de Creswell (2005).

Tabla 9.7 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 para elegir alcalde de Monterrey del 10 de noviembre, ¿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?
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 para elegir alcalde de Cortázar del 10 de noviembre, ¿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?

(continúa)

Tabla 9.7 Ejemplos de algunos problemas al elaborar preguntas (*continuación*)

Problema	Ejemplo de pregunta problemática	Mejora a la pregunta
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? ☐ Insatisfecho ☐ Medianamente satisfecho ☐ Satisfecho ☐ Sumamente satisfecho	Proporcionar equilibrio entre opciones favorables y desfavorables: ¿En qué medida está usted satisfecho con su superior inmediato? ☐ Sumamente insatisfecho ☐ Más bien insatisfecho ☐ Ni insatisfecho, ni satisfecho ☐ Más bien satisfecho ☐ Sumamente satisfecho
Incongruencia entre la pregunta y las opciones de respuesta	¿En qué medida está usted satisfecho con su superior inmediato? ☐ Muy poco importante ☐ Poco importante ☐ Medianamente importante ☐ Importante ☐ Muy importante	Generar categorías que coincidan con la pregunta: ¿En qué medida está usted satisfecho con su superior inmediato? ☐ Sumamente insatisfecho ☐ Más bien insatisfecho ☐ Ni insatisfecho, ni satisfecho ☐ Más bien satisfecho ☐ Sumamente satisfecho
Sólo una parte de los sujetos 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 que 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?
La pregunta puede ser inadecuada para parte de la población	¿Cómo le afectó el incremento en la tasa impositiva para empleados gubernamentales?	Agregar preguntas que segmenten a la población: ¿Actualmente trabaja? ☐ Sí ☐ No ¿Trabaja usted en... ☐ Empresa ☐ Por cuenta propia (independiente) ☐ Gobierno Entonces, a quienes pertenezcan a la última categoría, se les pregunta: ¿Cómo le afectó el incremento en la tasa impositiva para empleados gubernamentales?

Con relación a 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 sujeto 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 tal como: ¿Está de acuerdo con que se legalice el aborto en este país? Sin lugar a dudas sería un fracaso. Bostwick y Kyte (2005) señalan que los primeros cuestionamientos deben resultar interesantes para los sujetos. 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 preguntas demográficas sobre el estado civil, el género, la edad, la ocupación, el nivel de ingresos, el nivel educativo, la religión, la ideología, el puesto en una organización o algún tipo de afiliación a grupo, el partido y la institución. Pero en otras ocasiones es mucho mejor hacer este tipo de preguntas al final del cuestionario, particularmente cuando 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. Gambara (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 (Mertens, 2005).

¿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 sujeto, 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 (León y Montero, 2003).

A continuación, se presentan dos ejemplos sencillos de introducción y uno de instrucciones generales.

EJEMPLOS

Cartas introductorias

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.

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 sola opción; otras son de varias opciones y también se incluyen preguntas abiertas.

Muchas gracias por tu colaboración

Buenos días (tardes):

Comunicometría realiza una encuesta con el propósito de conocer las opiniones que se tienen acerca de esta empresa. Para ello le pediría fuera tan amable de contestar unas preguntas. No le tomará más de 20 minutos. La información que nos proporcione es muy importante para tomar decisiones y, desde luego, será manejada con la más estricta confidencialidad. No hay preguntas delicadas.

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.

Marque con claridad la opción elegida con una cruz o una "paloma". Recuerde: NO se deben marcar dos opciones. Marque así:

X ✓

Si NO PUEDE contestar una pregunta o si la pregunta no tiene sentido para usted, por favor pregúntele a la persona que le entregó este cuestionario y le explicó la importancia de su participación.

CONFIDENCIALIDAD

Sus respuestas serán anónimas y absolutamente confidenciales. Los cuestionarios serán procesados por personas externas. Además, como usted puede ver, en ningún momento se le pide su nombre.

De antemano: ¡MUCHAS GRACIAS POR SU COLABORACIÓN!

Algunas apelaciones que podemos utilizar en la introducción se muestran en la tabla 9.8.¹⁴

También se insertan instrucciones a lo largo del cuestionario. Las cuales nos indican cómo contestar. Por ejemplo:

—¿Tiene este ejido o esta comunidad ganado, aves o colmenas que sean de propiedad colectiva? (Marque con una cruz la respuesta.)

☐ Sí
(continúe)

☐ No
(pase a pregunta 30)

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

☐ 1 Sí
(pase a la pregunta 26)

☐ 2 No
(pase a la pregunta 27)

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

¹⁴ El concepto del recuadro está basado en Mertens (2005).

Tabla 9.8 Ejemplos de apelaciones para incentivar la participación

Apelación	Ejemplo
Incentivo	"Al responder, usted recibirá..." (dinero, un obsequio, un boleto, etcétera).
Altruismo	"Los resultados servirán para resolver...", "el estudio ayudará a..." (problema social, mejora en la calidad de vida, solventar una necesidad comunitaria, etcétera).
Autoconcepto de la persona	"Usted es una de las pocas personas que puede señalar ciertas cuestiones...", "debido a su experiencia (experticia, importancia, conocimientos, etc.) Usted puede... y por ello le solicitamos..." (su opinión calificada, etcétera).
Interés por el conocimiento	"Le enviaremos una copia de los resultados..."
Intereses profesionales	"Los resultados serán útiles para conocer temas importantes en nuestra profesión..."
Ayuda-auxilio	"Necesitamos su apoyo para conocer...", "los jóvenes requieren de ayuda para..."
Autoridad	Introducción acompañada de la firma de un líder o persona reconocida. O bien: "Doña Pola Castelan nos ha pedido que hagamos esta encuesta para conocer el problema de los niños...", "el científico...", "el empresario..."
Agradecimiento	"La comunidad de... estará muy agradecida por..."

EJEMPLO

Modelo de tarjeta de respuestas

	Completamente	Bastante	Regular	Poco	Nada
	(5)	(4)	(3)	(2)	(1)
• ¿Las políticas de su proveedor?					
• ¿Sus finanzas (estados financieros)?					
• ¿Los objetivos de su área de venta?					
• ¿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

¿ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ?	☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i>
¿ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ?	☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i>
¿ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ?	☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i>
¿ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ?	☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i>
¿ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ?	☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i> ☐ <i>~~~~~</i> <i>~~~~~</i> <i>~~~~~</i>

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

EJEMPLO

Modelo de formato horizontal

Presentación			
Preguntas sobre motivación intrínseca:			
¿ <i>~~~~~</i>	<i>~~~~~</i>	<i>~~~~~</i>	<i>~~~~~</i> ?
☐	☐	☐	☐
¿ <i>~~~~~</i>	<i>~~~~~</i>	<i>~~~~~</i>	<i>~~~~~</i> ?
☐	☐	☐	☐
Preguntas sobre satisfacción laboral			
Etcétera			

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 se pueden elaborar cuestionarios en CD y diseños de éstos para páginas web 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 sujetos 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 darles 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, basado parcialmente en Rojas (2002):

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

Decodificación 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
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, tal como se ejemplifica en el recuadro.

EJEMPLO

Categorías (frecuencias)

- Involucrando al personal y comunicándose con él (28)
 - Motivación e integración/mejoramiento del clima laboral (22)
 - Capacitación (14)
 - Incentivos/recompensas (11)
 - Difundiendo el valor "calidad" o la filosofía de la empresa (7)
 - Grupos o sesiones de trabajo (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 autoadministración puede tener distintos contextos: *individual, grupal o por envío (correo tradicional, correo electrónico y página web)*.

En el caso individual, el cuestionario se entrega al sujeto 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 sujetos 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 9.9 una lista de verificación de los aspectos centrales para administrar cuestionarios en grupo.

Tabla 9.9 Listado de puntos a verificar al administrar cuestionarios en grupo¹⁵

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 les notificó formalmente a los participantes potenciales la fecha, hora y 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 _____
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 lo suficiente fuerte para que todos la escuchen y su lectura será pausada?	Sí _____ Persona(s): _____	No _____ _____
10. ¿Se verificará que todos hayan respondido al cuestionario?	Sí _____	No _____
11. ¿Quién les 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 _____ _____

¹⁵ Adaptado de McMurthy (2005).

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 les 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 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 crítico en estos casos, ya que las posibilidades de retroalimentación 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. Ofrecerles a los participantes un resumen de los resultados, una vez que concluya la investigación, es una práctica recomendable (la cual se puede mandar por correo electrónico).

Asimismo, el paquete enviado a cada sujeto 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 uno 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 son elaborados en un programa de texto (procesador de palabras) o escaneados (si están impresos con anterioridad) y "anexados" en el correo electrónico (como un "archivo adjunto"), también colocarlo o "subirlo" al sitio web, aunque para este segundo caso lo más común es que se elaboren especialmente para tal ambiente. En ambos casos, 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) que 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 acceden 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, 2001) realizaron un par de estudios que se centraron en la utilización de la 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 (León y Montero, 2003). 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 (Vinuesa, 2005).

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).

2. Por entrevista personal

Las *entrevistas* implican que una persona calificada (entrevistador) aplica el cuestionario a los sujetos participantes, el primero hace las preguntas a cada sujeto 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 perturban, 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 sujetos que no hay respuestas correctas o equivocadas (Corbetta, 2003). Por otra parte, su proceder debe ser lo más estándar que sea posible (instrucciones iguales, 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.

Con relación al 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 sujeto entrevistado. 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 posean facilidad de palabra y capacidad de socializar. Como ya se explicitó previamente, los estudiantes de ambos géneros funcionan mucho mejor. Por lo tanto, es claro que para este fin deben rechazarse personas inseguras o excesivamente tímidas.

Rogers y Bouey (2005) diferencian entre la entrevista cuantitativa y la cualitativa, con relación a 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 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 siguiente apartado: Proceso de la investigación cualitativa.

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 un centro comercial, un parque o una escuela, etc.). Por ejemplo, Jaffe, 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 compra es menos costosa. En cualquier caso se aconseja que se busque un lugar lo más discreto, silencioso y privado que sea posible.

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.¹⁶

EJEMPLO

¿Cuál considera que es el principal problema en este municipio?

- | | | |
|--|---|--|
| ☐ Pobreza | ☐ Corrupción de funcionarios de la alcaldía | ☐ Desempleo |
| ☐ Falta de vivienda | ☐ Falta de infraestructura (calles, puentes, etcétera) | ☐ Empleo mal remunerado |
| ☐ Inseguridad en las calles | ☐ Problemas en la recolección de basura | ☐ Agua (escasez) |
-

También hace algunos años se ha generado 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 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), enton-

¹⁶ 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.

ces se le presenta al sujeto 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).

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. 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 sujeto 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 telemercadeo, 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:¹⁷

1. Enviar una carta previa, en la 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.
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.

¹⁷ Ciertas recomendaciones fueron extraídas de Mertens (2005).

4. Entre la carta previa y la comunicación 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).

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 (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.

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. Como en el siguiente ejemplo.

EJEMPLO



En desacuerdo



Neutral



De acuerdo

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 sujetos participantes (edad, nivel educativo, etcétera).

A continuación incluimos la tabla 9.10 que compara de manera sencilla las formas de administración.

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.

Tabla 9.10 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

Escalas para medir las actitudes

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; Oskamp, 1991; Eagly y Chaiken, 1993). 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 las primeras dos, que son las utilizadas 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 Likert¹⁸

Este método fue desarrollado por Rensis Likert en 1932; sin embargo, se trata de un enfoque vigente y bastante popularizado. Consiste en un conjunto de ítems presentados en forma de afirmaciones o juicios, ante los cuales se pide la reacción de los 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, BMW...), 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.

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.

¹⁸ 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).

EJEMPLO

Objeto de actitud medido	Afirmación
El voto	"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 afirmación correspondiente.¹⁹ Las opciones más comunes se presentan en la figura 9.13. Debe recordarse que a cada una de ellas se le asigna un valor numérico y sólo puede marcarse una respuesta. Se considera un dato inválido si se marcan dos o más opciones.

"Afirmación"

Alternativa 1:

Muy de acuerdo	De acuerdo	Ni de acuerdo, ni en desacuerdo	En desacuerdo	Muy en desacuerdo
----------------	------------	------------------------------------	---------------	----------------------

"Afirmación"

Alternativa 2:

Totalmente de acuerdo	De acuerdo	Neutral	En desacuerdo	Totalmente en desacuerdo
--------------------------	------------	---------	---------------	-----------------------------

"Afirmación"

Alternativa 3:

Siempre	La mayoría de las veces sí	Algunas veces sí, algunas veces no	La mayoría de las veces no	Nunca
---------	-------------------------------	---------------------------------------	-------------------------------	-------

"Afirmación"

Alternativa 4:

Completamente verdadero	Verdadero	Ni falso, ni verdadero	Falso	Completamente falso
----------------------------	-----------	---------------------------	-------	------------------------

Figura 9.13 Opciones o puntos en las escalas Likert.

Las opciones de respuesta o categorías pueden colocarse de manera horizontal, como en la figura 9.13, o en forma vertical.

() Muy de acuerdo

() De acuerdo

¹⁹ Likert (1932), Futrell *et al.* (1998), Clark (2000).

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

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.

Tabla 9.11 Opciones jerárquicamente correctas e incorrectas en un ejemplo²⁰

Objeto de actitud: Mi novia	
Correcto	Incorrecto (no se respeta la misma jerarquía en todos los ítems)
"Me gusta estar mucho con mi novia" ☐ Definitivamente sí ☐ Probablemente sí ☐ Indeciso ☐ Probablemente no ☐ Definitivamente no	"Me gusta estar mucho con mi novia" ☐ Probablemente sí ☐ Indeciso ☐ Definitivamente sí ☐ Probablemente no ☐ Definitivamente no
"Si por mí fuera, todos los días estaría con mi novia" ☐ Definitivamente sí ☐ Probablemente sí ☐ Indeciso ☐ Probablemente no ☐ Definitivamente no	"Si por mí fuera, todos los días estaría con mi novia" ☐ Definitivamente sí ☐ Probablemente sí ☐ Probablemente no ☐ Definitivamente no ☐ Indeciso
"Amo demasiado a mi novia" ☐ Definitivamente sí ☐ Probablemente sí ☐ Indeciso ☐ Probablemente no ☐ Definitivamente no	"Amo demasiado a mi novia" ☐ Definitivamente sí ☐ Probablemente sí ☐ Indeciso ☐ Probablemente no ☐ Definitivamente no

²⁰ Es un ejemplo simple para ilustrar el concepto.

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 afirmación estén los participantes, su actitud será igualmente más favorable.

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 lo tanto, *cuando las afirmaciones son positivas se califican comúnmente de la siguiente manera:*

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

Es decir, 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 afirmación, significa que su actitud es menos favorable, esto es, más desfavorable.

EJEMPLO

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

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 la afirmación, 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 9.14 se presenta un ejemplo de una escala Likert para medir la actitud hacia un organismo tributario.²¹

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	
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 9.14 Muestra de una escala Likert.

²¹ 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.

Como puede observarse en la figura 9.14, 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 9.15, la cual se basa en la figura 9.14, constituiría un ejemplo de cómo calificar una escala Likert.

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 po-

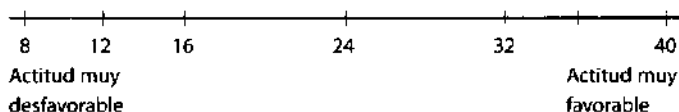
- | | |
|---|--|
| 1. El personal de la Dirección General de Impuestos Nacionales es grosero al atender al público. | |
| ☒ 1. Muy de acuerdo | ☐ 4. En desacuerdo |
| ☐ 2. De acuerdo | ☐ 5. Muy en desacuerdo |
| ☐ 3. Ni de acuerdo, ni en desacuerdo | |
| 2. La Dirección General de Impuestos Nacionales 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 | ☒ 3. 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 | |
| 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 | ☒ 3. 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 | ☒ 3. 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 | |

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

Figura 9.15 Muestra de puntuaciones de la escala Likert.

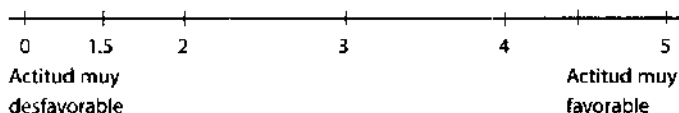
sible es de ocho ($1 + 1 + 1 + 1 + 1 + 1 + 1 + 1$) y la máxima es de 40 ($5 + 5 + 5 + 5 + 5 + 5 + 5 + 5$), porque hay ocho afirmaciones. La persona del ejemplo obtuvo 12. Su actitud hacia el organismo tributario es más bien bastante desfavorable; veámoslo gráficamente:

EJEMPLO



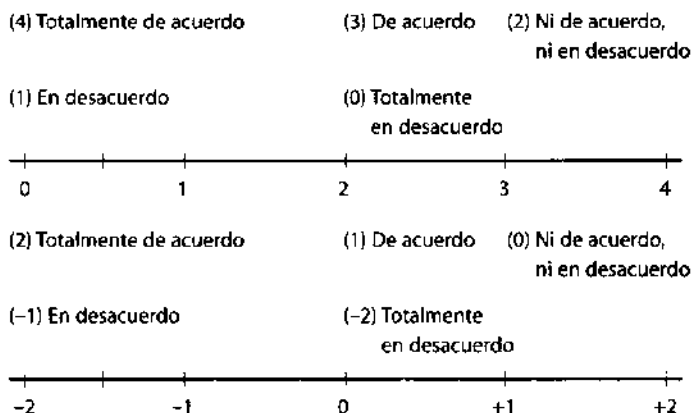
Si alguien hubiera tenido una puntuación de 37 ($5 + 5 + 4 + 5 + 5 + 4 + 4 + 5$) su actitud se calificaría como sumamente favorable. En las *escalas Likert* a veces se califica el promedio resultante en la escala mediante la sencilla fórmula PT/NT (donde PT es la puntuación total en la escala y NT es 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) señala que debe considerarse en un nivel de medición por intervalos porque ha sido probada en múltiples ocasiones. 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 sujetos 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 categorías. Pero debe recalarse 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 ítems.

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.

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 afirmación. Estas puntuaciones se correlacionan con las del grupo a toda la escala (la suma de las puntuaciones de todas las afirmaciones), y las afirmaciones, cuyas puntuaciones se correlacionen significativamente con las puntuaciones de toda la escala, se seleccionan para integrar el instrumento de medición. Asimismo, debe calcularse la confiabilidad y validez de la escala.

Preguntas en lugar de afirmaciones

En la actualidad, la escala original se ha extendido a preguntas y observaciones. Como se puede observar en el siguiente ejemplo.

EJEMPLO

¿Cómo considera usted al conductor que aparece en los programas?

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

La pregunta anterior se hizo como parte de la evaluación de un video empresarial. Otro ejemplo sería una pregunta formulada 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 9.12.

Tabla 9.12 Ejemplo de la escala Likert aplicada a varias preguntas

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

(continúa)

Tabla 9.12 Ejemplo de la escala Likert aplicada a varias preguntas (*continuación*)

¿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 tiempo de trabajar con el proveedor?					
la entrega del producto en las condiciones acordadas?					
Calidad del producto?					

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?

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.

Maneras de aplicar la escala Likert

Existen dos formas básicas de aplicar una escala Likert. La primera es de manera *autoadministrada*: se le entrega la escala al participante y éste marca, respecto de cada afirmación, la categoría que mejor describe su reacción. Es decir, marcan su respuesta. La segunda forma es *la entrevista*, donde un entrevistador lee las afirmaciones y alternativas de respuesta al sujeto, y anota lo que éste conteste. Cuando se aplica por medio de una entrevista, es necesario que se le entregue al entrevistado una tarjeta donde se muestren las alternativas de respuesta o categorías. El siguiente es un ejemplo que se aplica a la pregunta de la tabla 9.12:

EJEMPLO

De tarjeta de respuestas

Indispensable	Sumamente importante	Medianamente importante	Poco importante	No se toma en cuenta
---------------	----------------------	-------------------------	-----------------	----------------------

Al construir una escala Likert, debemos asegurarnos de que las afirmaciones y alternativas de respuesta serán comprendidas por los sujetos a los que se les aplicará y que éstos tendrán la capacidad de discriminación requerida. Ello se evalúa cuidadosamente en la prueba piloto.

Diferencial semántico

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

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 sujeto califica al candidato "A" en términos de esta escala de adjetivos bipolares.

Osgood, Suci y Tannenbaum (1957) nos indican que, si el sujeto 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

Si el sujeto 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

²² Para profundizar en el diferencial semántico se recomienda consultar: Osgood, Suci y Tannenbaum (1957, 1976a y 1976b), así como Heise (1976).

Si el sujeto 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 ejemplos 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.

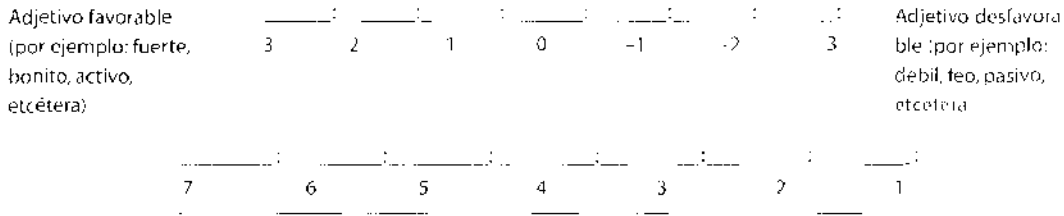
EJEMPLOS

Adjetivos bipolares

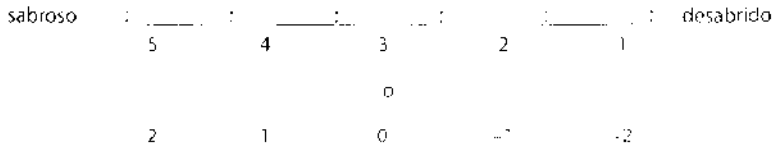
fuerte-débil	poderoso-impotente
grande-pequeño	vivo-muerto
bonito-feo	joven-vejez
alto-bajo	rápido-lento
claro-oscuro	gigante-enano
caliente-frío	perfecto-imperfecto
costoso-barato	agradable-desagradable
activo-pasivo	bendito-maldito
seguro-peligroso	arriba-abajo
bueno-malo	útil-inútil
dulce-amargo	favorable-desfavorable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

Codificación de las escalas

Los puntos o las categorías de la escala pueden codificarse de diversos modos, que se presentan en la figura 9.16.



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):

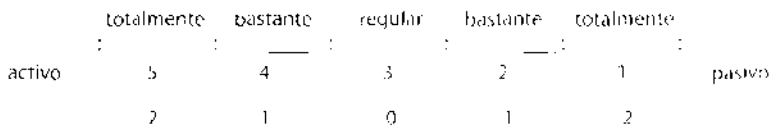
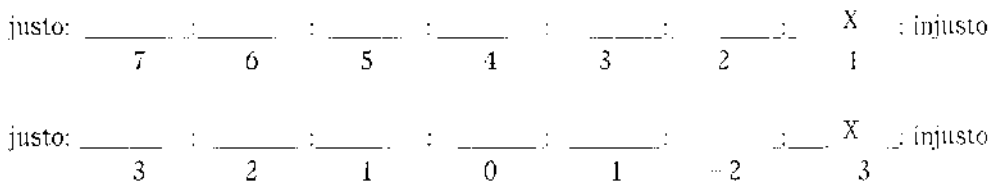


Figura 9.16 Maneras comunes de codificar el diferencial semántico.

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 calificación sería "1" o "-3".



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

Análisis de contenido

¿Qué es y para qué sirve el análisis cuantitativo de contenido?

Este método tiene mayores implicaciones que simplemente recoger contenidos, es al mismo tiempo un cierto tipo de diseño de investigación y/o técnica de análisis, pero lo incluimos en este capítulo porque innegablemente su proceso inicia con la recopilación de contenidos.

De acuerdo con la definición clásica de Berelson (1971), el **análisis de contenido** es una técnica para estudiar y analizar la comunicación de una manera objetiva, sistemática y cuantitativa. Krippendorff (1980) extiende la definición del *análisis de contenido* a un método de investigación para hacer inferencias válidas y confiables de datos con respecto a su contexto.

Resulta una técnica muy útil para analizar los procesos de comunicación en muy diversos contextos. El análisis de contenido puede aplicarse virtualmente a cualquier forma de comunicación (programas televisivos o radiofónicos, artículos en prensa, libros, poemas, conversaciones, pinturas, discursos, cartas, melodías, leyes y reglamentos, etc.). Por ejemplo, es conveniente para analizar la personalidad de alguien, evaluar sus escritos; conocer las actitudes de un grupo de personas mediante la evaluación de sus expresiones verbales o escritas; indagar sobre las preocupaciones de un pintor o un músico al estudiar su trabajo o material; compenetrarse con los valores de una cultura; o averiguar las intenciones de un publicista o propagandista.

Análisis de contenido

Se usa para estudiar y analizar la comunicación de una manera objetiva, sistemática y cuantitativa. Se basa en la definición de Berelson (1971) y Krippendorff (1980).

Usos del análisis de contenido

Berelson (1971) señala varios usos del análisis de contenido, entre los que destacan:

1. *Describir tendencias* en el contenido de la comunicación.
2. *Develar diferencias* en el contenido de la comunicación (entre personas, grupos, instituciones, países, etcétera).
3. *Comparar mensajes*, niveles y medios de comunicación.
4. *Auditar el contenido* de la comunicación y compararlo contra estándares u objetivos.
5. *Construir y aplicar estándares de comunicación* (políticas, normas, etcétera).
6. *Exponer técnicas publicitarias y de propaganda*.
7. *Medir la claridad* de los mensajes.
8. *Descubrir estilos* de comunicación.
9. *Identificar intenciones, apelaciones y características de comunicadores*.
10. *Descifrar mensajes ocultos* y otras aplicaciones a la inteligencia militar y a la seguridad política.
11. *Revelar "centros" de interés y atención* para una persona, un grupo o una comunidad.
12. *Determinar el estado psicológico* de personas o grupos.

13. *Obtener indicios del desarrollo verbal* (por ejemplo, en la escuela, como resultado de la capacitación, el aprendizaje de conceptos).
14. *Anticipar respuestas* a comunicaciones.
15. *Reflejar actitudes, valores y creencias de personas, grupos o comunidades.*
16. *Cerrar preguntas abiertas.*

El análisis de contenido se utiliza, digamos, para evaluar si varias series de televisión difieren entre sí en cuanto a su carga de contenido sexual; para conocer las discrepancias ideológicas entre varios periódicos o diarios (en términos generales o en referencia a un tema en particular); para comparar estrategias propagandísticas de partidos políticos; para contrastar, por medio de sus escritos, a distintos individuos que asisten a psicoterapia; para cotejar el vocabulario aprendido por pequeños que se exponen más al uso de la computadora en comparación con niños que la utilizan menos; para analizar la evolución de las estrategias publicitarias a través de algún medio en relación con un producto (como perfumes femeninos de costo elevado); para conocer y comparar la posición de diversos presidentes latinoamericanos en cuanto al problema de la deuda externa; para comparar estilos de escritores que se señalan como parte de una misma corriente literaria; etcétera.

¿Cómo se realiza el análisis de contenido?

El análisis de contenido se efectúa por medio de la *codificación*, es decir, el proceso en virtud del cual las características relevantes del contenido de un mensaje se transforman a unidades que permitan su descripción y análisis precisos. Lo importante del mensaje se convierte en algo susceptible de describir y analizar. Para codificar es necesario definir el *universo*, las *unidades de análisis* y las *categorías de análisis*.

Universo

El **universo** podría ser la obra completa de Franz Kafka; las emisiones de un noticiario o informativo en televisión durante un mes; los editoriales publicados en un día por cinco diarios o periódicos (por ejemplo, en España, *El Mundo*, *El País*, *Abc*, *La Razón*, *El Confidencial*); todos los capítulos de tres telenovelas; las canciones completas de un grupo o solista (U2, Joan Manuel Serrat, La Ley, Maná, etc.); los escritos de un conjunto de estudiantes durante un ciclo escolar; los discursos pronunciados por varios contendientes políticos, durante el último mes previo a la elección; los escritos de un grupo de pacientes en psicoterapia en un periodo específico; las conversaciones grabadas de 10 parejas que participan en un experimento sobre conflictos maritales; los escritos de un grupo terrorista, etc. El universo, como en cualquier investigación cuantitativa, debe delimitarse con precisión. Algunos autores como Sánchez Aranda (2005), se refieren al universo como el contexto de la unidad, el cuerpo más largo del contenido que puede examinarse al caracterizar una o más unidades de registro.

Unidades de análisis

Las **unidades de análisis** o **registro** constituyen segmentos del contenido de los mensajes que son caracterizados para ubicarlos dentro de las categorías. Sánchez Aranda (2005, p. 121), las define como: “el cuerpo de contenido más pequeño en que se cuenta la aparición de una referencia, ya sean palabras o afirmaciones que nos interesa localizar”. Berelson (1971) menciona cinco unidades importantes de análisis:

1. *La palabra.* Es la unidad de análisis más simple, aunque, como señala Kerlinger y Lee (2002), puede haber unidades más pequeñas como letras, fonemas, símbolos. Así, se puede medir cuantas veces aparece una palabra en un mensaje (por ejemplo, número de veces que en los informativos o noticiarios televisivos de fin de año se menciona al presidente de la República).
2. *El tema.* Se define a menudo como un enunciado respecto de algo. Los temas suelen ser más o menos generales. Kerlinger y Lee (2002) utilizan un excelente ejemplo para ello: “Las cartas de adolescentes o estudiantes de colegios superiores pueden ser estudiadas en sus expresiones de autorreferencia”. Éste sería un tema extenso, pero se definirían los temas como cualesquiera de las oraciones que usen “yo”, “mi” y otros términos que indiquen referencia al “yo” del escritor. Así, se analizaría que tanta autorreferencia contienen las cartas.

Si los temas son complejos, el análisis del contenido es más difícil, en especial si se complica al incluirse más de una oración simple.

El ítem. Tal vez es la unidad de registro más utilizada y puede definirse como la unidad total empleada por los productores del material simbólico (Berelson, 1971). Ejemplos de ítems son un libro, un editorial, un programa de radio o televisión, un discurso, un libro, una fotografía, una carta amorosa, una conversación telefónica, una canción o la respuesta a una pregunta abierta. En este caso, lo que se analiza es el material simbólico total.

3. *El personaje.* Se trata de un individuo. Por ejemplo, un paciente, un héroe de una serie televisiva, un líder histórico, etc. Aquí lo que se analiza es el personaje.
4. *Medidas de espacio-tiempo.* Son unidades físicas como el centímetro columna (por ejemplo, en la prensa), la línea (en escrito), el minuto (en una conversación o en radio), el periodo de 20 minutos (en una interacción), el cuadro (en televisión), cada vez que se haga una pausa (en un discurso). Sánchez Aranda (2005) le denomina unidad de enumeración. En este caso, la unidad física (por ejemplo, cada minuto) es ubicado en categorías (si tenemos 10 minutos, diez unidades son colocadas en las categorías).

Las unidades se insertan, colocan o caracterizan en categorías y/o subcategorías; esto podría representarse tal como se muestra en la figura 9.19.

Al cuestionarnos sobre qué unidad debe seleccionarse, esto depende de los objetivos y las preguntas de investigación. Sin embargo, Berelson (1971) sugiere lo siguiente:

- a) En un solo estudio es posible utilizar *mas de una unidad de análisis*.
- b) Los cálculos de palabras y las unidades amplias, como el ítem y las medidas de espacio tiempo, son más adecuados en los análisis que dan mayor importancia a los asuntos definidos.
- c) Las *unidades* amplias y las más definidas son válidas para la aceptación o el rechazo en una categoría.

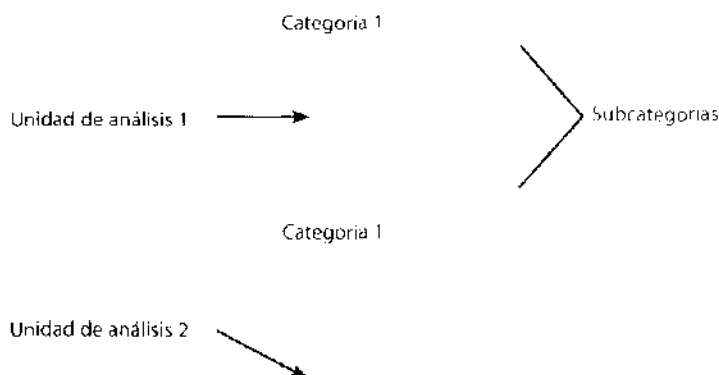


FIGURA 9.19 Agrupación de unidades de análisis en categorías.

- d) Las *unidades* amplias generalmente requieren de menos tiempo para su codificación que las unidades pequeñas, referidas a las mismas categorías y materias.
- e) Debido a que los temas o las oraciones agregan otra dimensión al asunto, la mayoría de las veces son más difíciles de analizar que las palabras y las *unidades* amplias.
- f) El tema es adecuado para el análisis de significados y las relaciones entre éstos.

Categorías

Las **categorías** son los niveles donde serán caracterizadas las unidades de análisis. Como menciona Holsti (1969), son las “casillas o cajones” en las cuales se clasifican las unidades de análisis. Sánchez Aranda (2005) las define como características o aspectos que presenta la comunicación con la que estamos trabajando (en cuanto a referencias). Por ejemplo, un discurso podría clasificarse como optimista o pesimista, como liberal o conservador. Un personaje de una caricatura se clasificaría como bueno, neutral o malo. En ambos casos, la unidad de análisis se categoriza. Veámoslo esquemáticamente en la figura 9.20.

Es decir, cada unidad de análisis se categoriza o encasilla en uno o más sistemas de categorías. Por ejemplo, en un estudio citado por Krippendorff (1980) se analizaron 2 430 actos de violencia televisada. En cada acto, el personaje principal (unidad de análisis) se categorizó como:

- Bueno, neutral o malo (sistema 1).
- Asociado con hacer cumplir la ley, no relacionado con el cumplimiento de la ley o se presentaba como un delincuente o criminal (sistema 2).

En este caso, la unidad de análisis o registro es el comportamiento del personaje durante el acto televisivo, y las *categorías* eran dos: bondad del personaje y carácter involucrado. A su vez, las subcategorías de la bondad del personaje eran tres: bueno, neutral y malo. Y las subcategorías del carácter involucrado también eran tres: asociado con cumplir la ley, no relacionado con la ley y un criminal. Esto podría representarse como se muestra en la figura 9.21.

La selección de categorías también depende del planteamiento del problema.

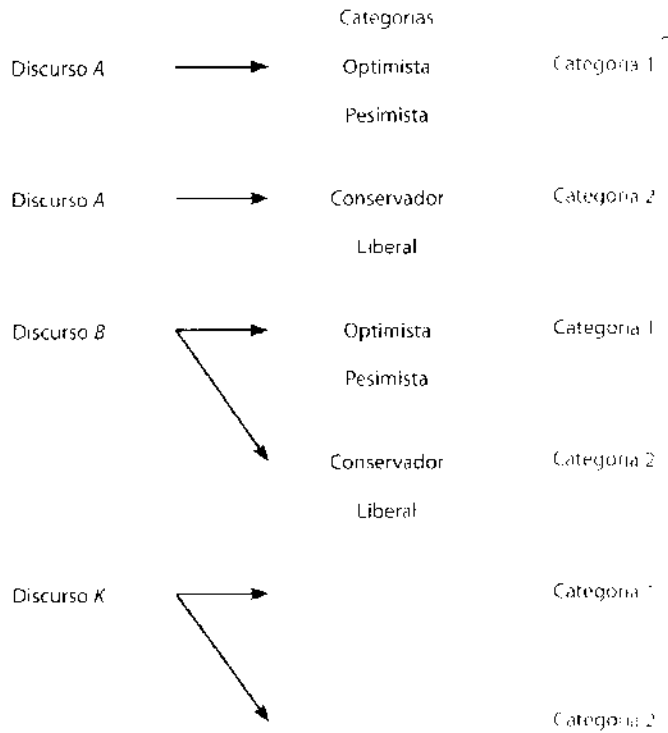


Figura 9.20 Ejemplos de unidades de análisis y su clasificación en categorías.

Unidades	Categorías	Subcategorías
Acto 1	Bondad del personaje	{ bueno neutral malo }
Acto 2		
Acto 3		
Acto 4		
•		
•	-Caracter involucrado	{ asociado con cumplir la ley no relacionado con la ley criminal }
Acto 2430		

Figura 9.21 Ejemplo con dos categorías y seis subcategorías.

Tipos de categorías

Krippendorff (1980) señala cinco tipos de categorías:

1. *De asunto o tópico.* Se refieren a cuál es el asunto, materia o temática tratada en el contenido (¿de qué trata el mensaje o la comunicación?).

EJEMPLO

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

Categoría:	Tema financiero
Subcategorías:	Deuda
	Impuestos
	Planeación hacendaria
	Inflación
	Etcétera

2. *De dirección.* Estas categorías se refieren a cómo es tratado el asunto (¿positiva o negativamente?, ¿favorable o desfavorablemente?, ¿nacionalista o no nacionalista?, etcétera).

EJEMPLO

Comparar la manera en que dos noticiarios televisivos hablan de la posibilidad de una moratoria unilateral en el pago de la deuda externa de América Latina.

Categoría:	Tono en el tratamiento de la deuda externa.
Subcategorías:	A favor de la moratoria unilateral
	En contra
	Etcétera

3. *De valores.* Este tipo de categorías indican qué valores, intereses, metas, deseos o creencias son revelados.

EJEMPLO

Al estudiar la compatibilidad ideológica de matrimonios, se podría analizar la ideología de cada cónyuge, pidiéndoles un escrito sobre temas que puedan reflejar valores (sexo, actitud hacia la pareja, significado del matrimonio).

Categoría:	Ideología del esposo
Subcategorías:	Muy tradicional
	Más bien tradicional
	Neutral
	Más bien liberal
	Muy liberal

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

Figura 9.21

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

Categoría:	Receptores a quienes se dirige el mensaje
Subcategorías:	Opinión pública en general
	Presidente
	Gabinete económico
	Gabinete agropecuario
	Gobierno en general
	Sector empresarial
	Obreros afiliados a su sindicato
	Obreros no afiliados a su sindicato
	Obreros en general (afiliados y no afiliados)
	Etcétera

5. *Físicas.* Son categorías para ubicar la posición y duración o extensión de una unidad de análisis. De *posición* son, por ejemplo, la sección y página (en prensa), y el horario (en televisión y radio); de *duración*, los minutos (en una interacción, un comercial televisivo, un programa de radio, un discurso), los centímetros/columna (en prensa), los cuadros en una película, etcétera.

Ejemplo 9.2

Cada periodo de 10 minutos de una telenovela se va a considerar la unidad de análisis.

							unidad
1	2	3	4	5	6	7	k
10 m	10 m	10 m	10 m	10 m	10 m	10 m	10 m

duración de la telenovela (categoría)

La unidad 1 la colocó en categorías.

La unidad 2 la colocó en categorías.

La unidad K la colocó en categorías.

No se deben confundir las unidades de espacio-tiempo con las categorías físicas. Las primeras son unidades de análisis y las segundas constituyen categorías. Veámoslo en la figura 9.22 con un par de ejemplos simples.

Tiempo como unidad

Analizar las referencias
a la ideología terrorista en
los discursos de cada
presidente.
Unidad de análisis:
Discurso

Unidad 1:
Unidad 2:
Unidad 3:
Unidad 4:
Etcétera



Categorías de análisis:

1) Referencias al grupo:

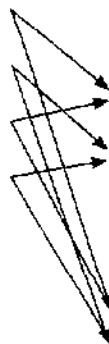
- Asesinos
- Homicidas
- Inhumanos
- Malévolos
- Crueles
- Injustos

Etcétera

Tiempo como categoría

Analizar las referencias
a la ideología terrorista en
los discursos de mandatos
del presidente o
gubernador de cada
estado de la Unión
Americana.

Unidad 1: Discurso 1
Unidad 2: Discurso 2
Unidad 3: Discurso 3
Unidad 4: Discurso 4
Etcétera



Categorías de análisis:

1) Referencias al grupo:

- Asesinos
- Homicidas
- Inhumanos
- Malévolos
- Crueles
- Injustos

Etcétera

2) Tiempo de duración del
discurso.

- 1 a 10 minutos
- 11 a 20 minutos
- 21 a 30 minutos
- 31 a 40 minutos

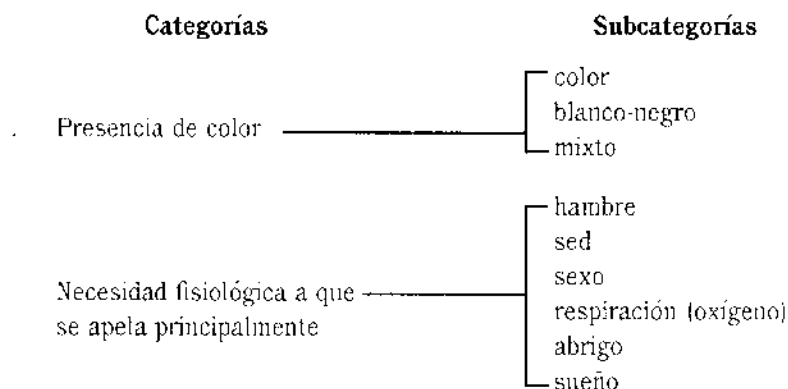
Etcétera

Ejemplo 4.2 Ejemplos del tiempo como unidad de análisis y como categoría.**Requisitos de las categorías.**

En un análisis de contenido se suelen tener varias categorías, que deben cumplir los siguientes requisitos:

1. Las categorías y subcategorías deben ser *exhaustivas*. Es decir, abarcar todas las posibles subcategorías de lo que se va a codificar. Por ejemplo, la categoría "ideología del esposo" no podría prescindir de la subcategoría "neutral".
2. Las subcategorías deben ser de preferencia *mutuamente excluyentes*, de tal manera que una unidad de análisis clasifique en una y sólo una de las subcategorías de cada categoría. Por ejemplo, un personaje no puede ser "bueno" y "malo" a la vez.

Obviamente, una unidad de registro encajará en una subcategoría de una categoría y en otras subcategorías de las demás categorías. Por ejemplo, al analizar comerciales televisivos podríamos tener, entre otras categorías, las siguientes:



Una unidad de análisis (un comercial) puede seleccionarse en una subcategoría de "presencia de color" y en otra subcategoría de la categoría "necesidad fisiológica a que se apela" (por ejemplo: color y sed); pero no seleccionarse en dos subcategorías de la misma categoría, como "hambre" y "sed" (porque la categoría es la necesidad a que se apela *principalmente*), a menos que generáramos la subcategoría "hambre y sed".

En ciertos casos especiales, llega a interesar al analista una categoría donde las subcategorías no sean mutuamente excluyentes. Tal sería el caso de que la categoría fuera "apelaciones del comercial" (estatus socioeconómico, salud, juventud, belleza, amor romántico, etc.). Un comercial podría utilizar más de una apelación (salud y belleza, por ejemplo).

3. Las categorías y subcategorías *deben derivarse del marco teórico y de una profunda evaluación de la situación.*

Ejemplo de un análisis de contenido

Para ejemplificar el análisis de contenido y específicamente la generación de categorías, se acudirá a un estudio de Greenberg, Edison, Korzeny, Fernández-Collado y Atkin (1980). El estudio consistió en un análisis de contenido de las series televisadas por las tres grandes cadenas de Estados Unidos: ABC, CBS y NBC. Se evaluaron diversos programas durante tres periodos de 1975 a 1978, para determinar el grado en que la televisión estadounidense mostraba actos prosociales y antisociales como medida de la violencia televisada. Las categorías y subcategorías utilizadas son las que se muestran en la figura 9.23.

El estudio consideró como unidad de análisis la conducta, cada vez que se presentaba una conducta de los personajes era codificada.

Cuando se crean las categorías, éstas deben definirse con precisión y es necesario explicitar qué se va a comprender en cada caso, y qué habrá de excluirse.

El *análisis de contenido* consiste en asignar cada unidad a una o más categorías. De hecho, el producto de la codificación son frecuencias de categorías. Se cuenta cuántas veces se repite cada categoría o subcategoría (cuántas unidades de análisis se clasificaron en cada una de las categorías). Por ejemplo, Greenberg *et al.* (1980, p. 113) encontraron los resultados que se muestran en la tabla 9.13.

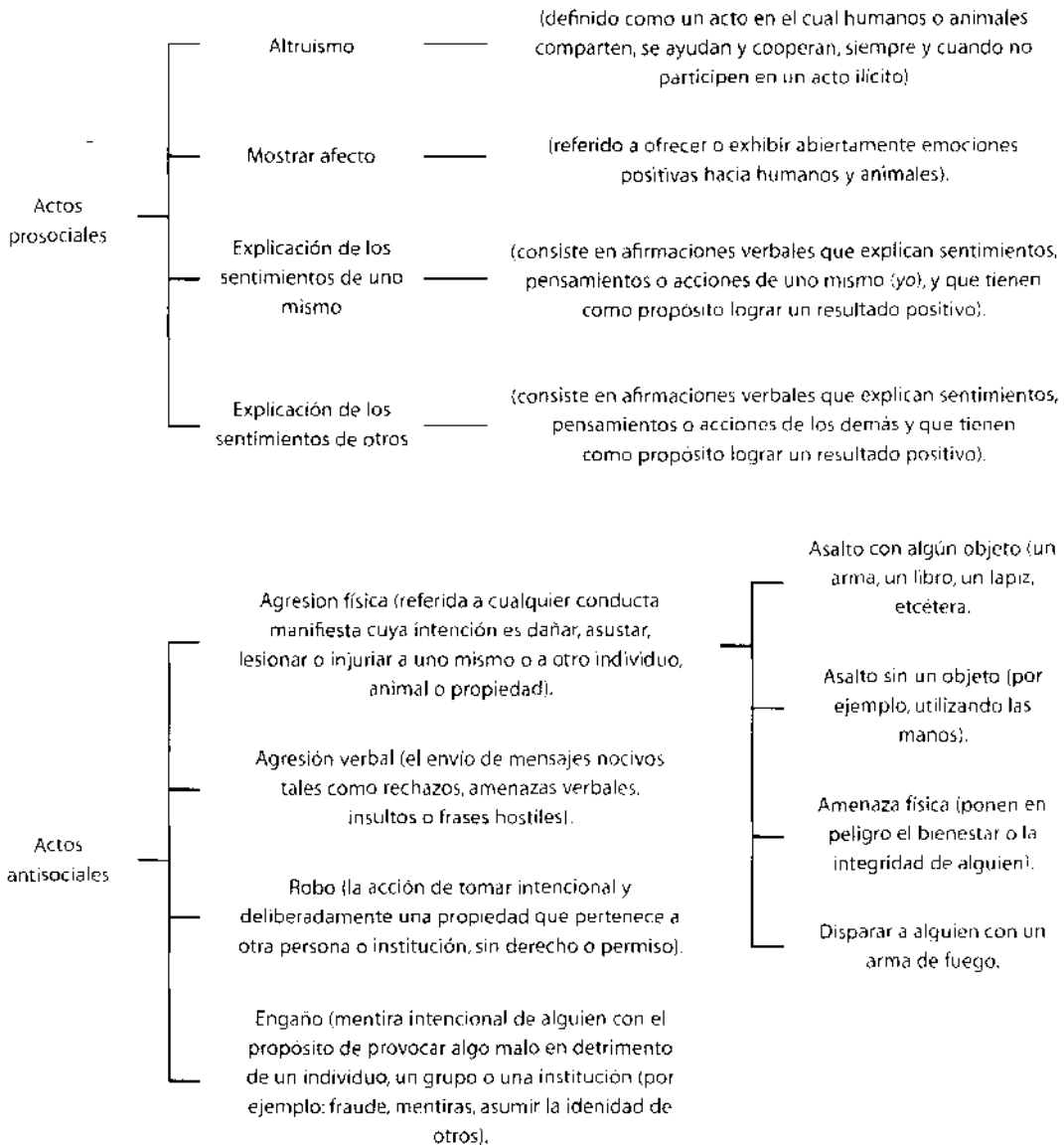


Figura 9.23 Ejemplo de sistema de categorías.

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

Incidencia de actos antisociales en los tres periodos

	Año 1*		Año 2*		Año 3*	
	f	%	f	%	f	%
I. Agresión física						
A. Asalto con un objeto	466	(15.7)	248	(10.8)	370	(13.6)
B. Asalto sin un objeto	111	(3.7)	159	(6.9)	177	(6.5)
C. Amenaza física	180	(6.1)	233	(10.1)	135	(5.0)
D. Disparar	106	(3.6)	75	(3.2)	74	(2.7)
E. Otras	128	(4.3)	171	(7.4)	130	(4.8)
II. Agresividad verbal	1629	(55.0)	1099	(47.6)	1464	(54.0)
III. Robo	61	(2.1)	72	(3.1)	44	(1.6)
IV. Engaño	283	(9.5)	251	(10.9)	319	(11.8)
Total	2964		2308	(100.0)	2713	(100.0)
Horas analizadas	(68.5)		(58)		(63)	

*Año 1 incluyó de octubre de 1975 hasta que se grabó un episodio de cada una de las series existentes (1976); año 2 igual, pero en 1976-1977, y año 3 igual, pero de 1977-1978.

f = número de casos o unidades.

% = porcentajes.

Ejercicio (inicio)

Imagine el lector que le asignan la tarea de clasificar películas u obras cinematográficas para evaluar si son adecuadas para el público de todas las edades, si solamente son apropiadas para adolescentes y adultos, o únicamente son aptas para adultos. Por supuesto, la unidad de análisis es el filme o película, pero: ¿qué categorías podrían utilizarse para clasificar a las unidades? Reflexione y más adelante continuamos con el ejercicio.

— Los pasos de la investigación cuantitativa —

Ya hemos mencionado tres:

Definir con precisión el universo y extraer una muestra representativa.

Establecer y definir las unidades de análisis

Establecer y definir las categorías y subcategorías que representan a las variables de la investigación.

Los demás pasos son:

Seleccionar los codificadores. Los codificadores son las personas que habrán de asignar las unidades de análisis a las categorías. Deben tener un nivel educativo profesional (estudiantes a nivel de licenciatura como mínimo).

Elaborar las hojas de codificación. Estas hojas contienen las categorías y subcategorías, y los codificadores anotan en ellas cada vez que una unidad entra en una subcategoría. En la figura 9.24 se muestran los elementos de una hoja de codificación.

Fecha en que se codificó el material →

Indicador de quien codificó: número, iniciales o letras →

Frecuencias (veces que se repite cada categoría) →

Descripción del material que se analizará (discurso, nota periodística, sesión terapéutica, etcétera) →

CODIFICADOR: 1
FECHA: 29-XI-02

CAPÍTULO DE LA SERIE "ALF" (CBS)

MATERIAL A ANALIZAR: DURACIÓN: 30 MINUTOS

FRECUENCIAS

TOTALES

Suma de frecuencias en cada categoría o subcategoría

Total de frecuencias o unidades

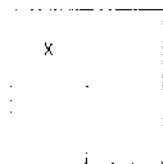
Categorías y subcategorías

A C T O S	ALTRUISMO (1.1.0)		2
	MOSTRAR AFECTO (1.2.0)		5
	EXPLICACIÓN DE LOS PROPIOS SENTIMIENTOS (1.3.0)		3
	EXPLICACIÓN DE LOS SENTIMIENTOS DE OTROS (1.4.0)		0
	AGRESIÓN FÍSICA 2.1		0
A C T O S	Asalto con objeto (2.1.1)		0
	Asalto sin objeto (2.1.2)		0
	Amenaza física (2.1.3)		0
	Disparar (2.1.4)		0
	Otros (2.1.5)		0
A N T I S O C I A L E S (2)	AGRESIÓN VERBAL (2.2.0)		3
	ROBO (2.3.0)		0
	ENGAÑO (2.4.0)		1
TOTAL		14	

Al reverso pueden solicitarse comentarios del codificador

Elementos de una hoja de codificación.

Las *hojas de codificación* pueden incluir elementos más precisos para describir el material. Por ejemplo, para analizar artículos de periódicos: nombre del periódico, fecha de publicación, género periodístico (reportaje, entrevista, editorial, artículo de fondo, carta, inserción pagada, artículo general), signatario del texto, sección donde se publicó, página, volumen de publicación (año, número o equivalente), ubicación en los cuadrantes de páginas.



tamaño (en centímetros/columna), nombre del codificador, día de codificación, etc. Una hoja puede servir para una o varias unidades, según nos interese o no tener datos específicos de cada unidad (normalmente se prefiere tener la información por unidad). Las hojas pueden ser impresas o estar archivadas en la computadora (en algún programa de texto o base de datos, por ejemplo: Word, Excel, etcétera).

Asimismo, las *categorías y subcategorías* deben *codificarse* con sus respectivos valores numéricos. En la hoja de codificación de la figura 9.24, 1 (uno) significa acto prosocial y 2 (dos), acto antisocial (así, 1.1.0 es "altruismo", 1.2.0 "mostrar afecto", etc.). Aparecen tres cifras en cada caso, porque hay categorías (actos prosociales y antisociales), subcategorías primarias (por ejemplo, explicación de los propios sentimientos, agresión física) y subcategorías secundarias (asalto con objeto, asalto sin objeto, amenaza física).

6. *Proporcionar entrenamiento a codificadores.* Esta capacitación incluye que los codificadores se familiaricen y compenentren con las variables, comprendan las categorías y subcategorías y entiendan las definiciones de ambas. Además, debe entrenárseles en la manera de codificar y discutir ampliamente con ellos las diferentes condiciones en que puede manifestarse o estar presente cada categoría y subcategoría. Asimismo, los codificadores tienen que comprender en forma cabal en qué consiste la unidad de análisis.

Parte del entrenamiento implica realizar una prueba piloto con el sistema de categorías y subcategorías, es decir, realizar un análisis de contenido de una muestra representativa del universo del material de estudio (por ejemplo, unos programas de las series televisivas a analizar, segmentos de los libros de un autor —capítulos de cada obra—, partes de los discursos de oradores —páginas de cada discurso—; en este caso, en lugar de muestrear con individuos, se realiza con unidades de significado). Así, podemos descubrir categorías y/o subcategorías no contempladas previamente. La prueba piloto también nos sirve para evaluar qué tan confiables son los codificadores, lo que nos conduce a la siguiente etapa.

7. *Calcular la confiabilidad de los codificadores.* Todos los codificadores realizan análisis de contenido provisional de una parte representativa del material (éste debe ser el mismo para todos los codificadores), con la finalidad de determinar el grado de acuerdo o consenso entre ellos. Si no hay consenso no puede efectuarse una evaluación confiable de la muestra o universo de contenido.

Para lo anterior se calcula la confiabilidad de cada codificador (individual) y la confiabilidad entre codificadores. El cálculo de la *confiabilidad* individual de los codificadores depende de si tenemos uno o varios de ellos.

a) Si se dispone de un solo codificador (porque el material es reducido) se observan las diferencias de la codificación del mismo mensaje hecha por el codificador en dos tiempos diferentes. Si las diferencias son muy pequeñas, el codificador es individualmente confiable. Este tipo de confiabilidad se llama *confiabilidad intracodificador*, la cual mide la estabilidad de la prueba y reprueba de un codificador a través del tiempo.

Otro método consistiría en que el codificador trabaje una parte representativa del material y después, un experto externo y el investigador, revisan su trabajo y aplican a su codificación (resultados) la siguiente fórmula:

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

Por ejemplo, un mensaje que conste de 20 unidades y se logren correctamente las 20, la confiabilidad será de $1\left(\frac{20}{20}\right)$, que es el máximo de confiabilidad. Si el codificador solo pudo codificar en forma correcta 15 de los 20, la confiabilidad sería de $0.75\left(\frac{15}{20}\right)$.

b) Si se dispone de varios codificadores, la confiabilidad individual puede determinarse así: se solicita a todos los codificadores que codifiquen el mismo material, se toman los resultados de todos menos los de uno y se compara la codificación de éste contra la del resto. Así se procede con cada codificador.

También puede aplicarse a todos los codificadores la fórmula mencionada para calcular la confiabilidad individual; quien se distancie del resto se considera un caso poco confiable.

EJEMPLO

Codificador A
0.89

Codificador B
0.93

Codificador C
0.92

Codificador D
0.67

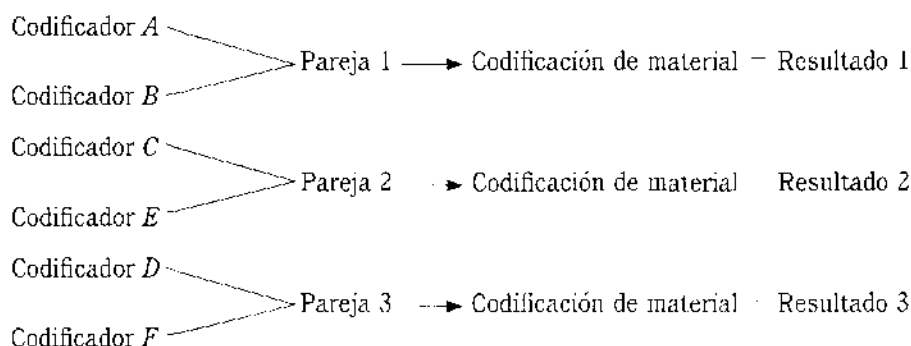
El codificador D tiene baja confiabilidad.

El cálculo de la *confiabilidad intercodificadores* se realiza por pares de codificadores (parejas). Se pide a cada pareja formada que codifique el material, se comparan los resultados obtenidos por las parejas, se cuenta el número de acuerdos entre las parejas, se determina el número de unidades de análisis y se aplica la siguiente fórmula:

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

Después se suman los resultados de esta fórmula y se divide entre el número de comparaciones, que depende del número de parejas. Veámoslo con un ejemplo:

Las parejas codifican el material:



Se determina el número de acuerdos entre las parejas (un acuerdo consiste en que dos parejas codifican en la misma categoría a una misma unidad de análisis).

<i>Pareja</i>	<i>Número de unidades de análisis codificadas</i>	<i>Número de acuerdos entre parejas</i>
1	18	Entre parejas 1 y 2 17
2	18	Entre parejas 1 y 3 16
3	17	Entre parejas 2 y 3 16

Se aplica la fórmula de confiabilidad entre parejas

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

$$C_{1,3} = \frac{16}{18} = 0.89$$

$$C_{2,3} = \frac{16}{18} = 0.89$$

Debe observarse que no hubo consenso total entre cuántas unidades de análisis podían distinguirse en el material [la pareja 3 distinguió 17 unidades, y las parejas 1 y 2 distinguieron 18]. En este caso, para la fórmula de confiabilidad entre parejas, se toma en cuenta el máximo de unidades de análisis codificadas por alguna de las parejas. Si fuera:

<i>Pareja</i>	<i>Número de unidades de análisis codificadas</i>	<i>Número de acuerdos entre parejas</i>
A	25	se toma el máximo 21
B	22	

La fórmula de confiabilidad para este último caso sería:

$$C_{AB} = \frac{21}{25} = 0.84$$

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

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

$$C_T = \frac{0.04 + 0.89 + 0.89}{3} = 0.90$$

No es conveniente tolerar una *confiabilidad* menor de 0.85 (ni total, ni entre dos parejas) y de ser posible debe superar el 0.80. Al igual que con otros instrumentos de medición, la confiabilidad oscila entre cero (0 = nula confiabilidad) y uno (1 = confiabilidad total).

En el análisis de contenido una confiabilidad baja puede deberse a que las categorías o unidades de análisis no se definieron con claridad y precisión, a un deficiente entrenamiento o a la inhabilidad de los codificadores. Cuando se obtiene una baja confiabilidad tiene que detectarse y corregirse el problema.

Asimismo, es conveniente calcular la confiabilidad a la mitad de la codificación (con el material codificado en un 50%) y al completar ésta.

Efectuar la codificación. Lo que implica asignar las unidades al sistema de categorías y contar las frecuencias de repetición de las unidades en las categorías (número de unidades que se seleccionan para cada categoría).

Ordinar los datos de las hojas de codificación y obtener totales de cada categoría.

Realizar los análisis estadísticos apropiados. Para esto último, debe revisarse el siguiente capítulo.

Ejercicio (continuación)

Algunas de las categorías para clasificar a las películas, cintas o filmes podrían ser:

- Contenido de sexo.
- Contenido de violencia.
- Contenido de lenguaje ofensivo o procaz (insultos, groserías, majaderías).
- Contenido sobre adicciones y drogas.
- Contenido de horror (no terror, sino por ejemplo escenas de muertos —un holocausto que muestre cadáveres— o la filmación de la muerte de una persona).

Cabrían al menos dos preguntas sobre estas categorías: ¿son demasiado generales?, ¿qué significa cada una de ellas?, ¿son suficientes o podrían establecerse otros criterios? Las respuestas podrían ir en varios sentidos. La primera es indudable: “sí”, tales categorías son demasiado generales. Por ello, su significado es poco preciso. Y alguien podría agregar otra categoría: contenido de discriminación, y aun otras categorías.

La categoría “contenido de sexo” podría englobar las siguientes subcategorías primarias: a) desnudos, b) presentación de actos sexuales, c) degradación de valores mediante el sexo,

y otros conceptos. En la subcategoría desnudos, podrían agruparse otras subcategorías (secundarias), por ejemplo: 1) eróticos ligados al acto sexual (uno o varios actores y actrices desnudos, comprometidos en acciones sexuales), 2) eróticos no ligados al acto sexual (uno o varios actores y actrices simplemente desnudos), 3) no eróticos ni ligados a una acción sexual (por ejemplo, una escena donde aparezca alguna escultura desnuda, como el David de Miguel Ángel o un bebé desnudo). La presentación de actos sexuales podría abarcar las siguientes subcategorías secundarias: 1) sexo sugerido (escenas donde los actores insinúan que habrá un acto sexual, pero no se ve ninguna acción; por ejemplo, entran a una recámara, sonríen, se miran entre sí, se cierra la puerta), 2) sexo implícito (escenas de conducta sexual manifiesta, en la que se observa una acción pero ésta no es clara ni precisa por ejemplo, una pareja que está en una cama, entre las sábanas y supuestamente tiene una relación sexual) y 3) sexo explícito (conducta sexual manifiesta y escenificada, en la cual los actores se encuentran desnudos y se observa con claridad y precisión la acción sexual). La degradación de valores podría clasificarse en (subcategorías secundarias): 1) el sexo se presenta como una forma de entrega y amor de la pareja para el desarrollo armónico de la relación y el crecimiento humano de sus integrantes, 2) el sexo se presenta como una pulsión instintiva, 3) el sexo se presenta como una forma de degradación humana o humillación.

La violencia podría ser de distintos tipos (subcategorías primarias): a) física, b) verbal o c) psicológica. En el caso de la violencia física, ésta podría llevarse a cabo con armas o sin ellas (subcategorías secundarias). La violencia física podría provocar ciertos efectos: 1) Ningún efecto, 2) heridas reversibles que no requieren hospitalización, 3) heridas reversibles que requieren hospitalización, 4) heridas irreversibles que requieren hospitalización y 5) muerte (subcategorías secundarias). La violencia también podría asociarse al consumo de drogas o no y al contenido sexual o no (subcategorías secundarias).

El contenido sobre adicciones y/o drogas, podría abarcar las subcategorías primarias: a) producción, b) distribución y c) consumo. Esta última podría subdividirse en subcategorías secundarias: 1) consumo de alcohol, 2) tabaco, 3) estupefacientes de daño significativo pero menor (por ejemplo, marihuana) y 4) estupefacientes de daño severo a todo el organismo humano (por ejemplo, cocaína, heroína, entre otras). Asimismo, otras subcategorías secundarias del consumo podrían ser en relación con sus consecuencias: 1) daños reversibles a los sistemas del organismo, 2) daños reversibles y 3) fallecimiento. Y así, las demás categorías.

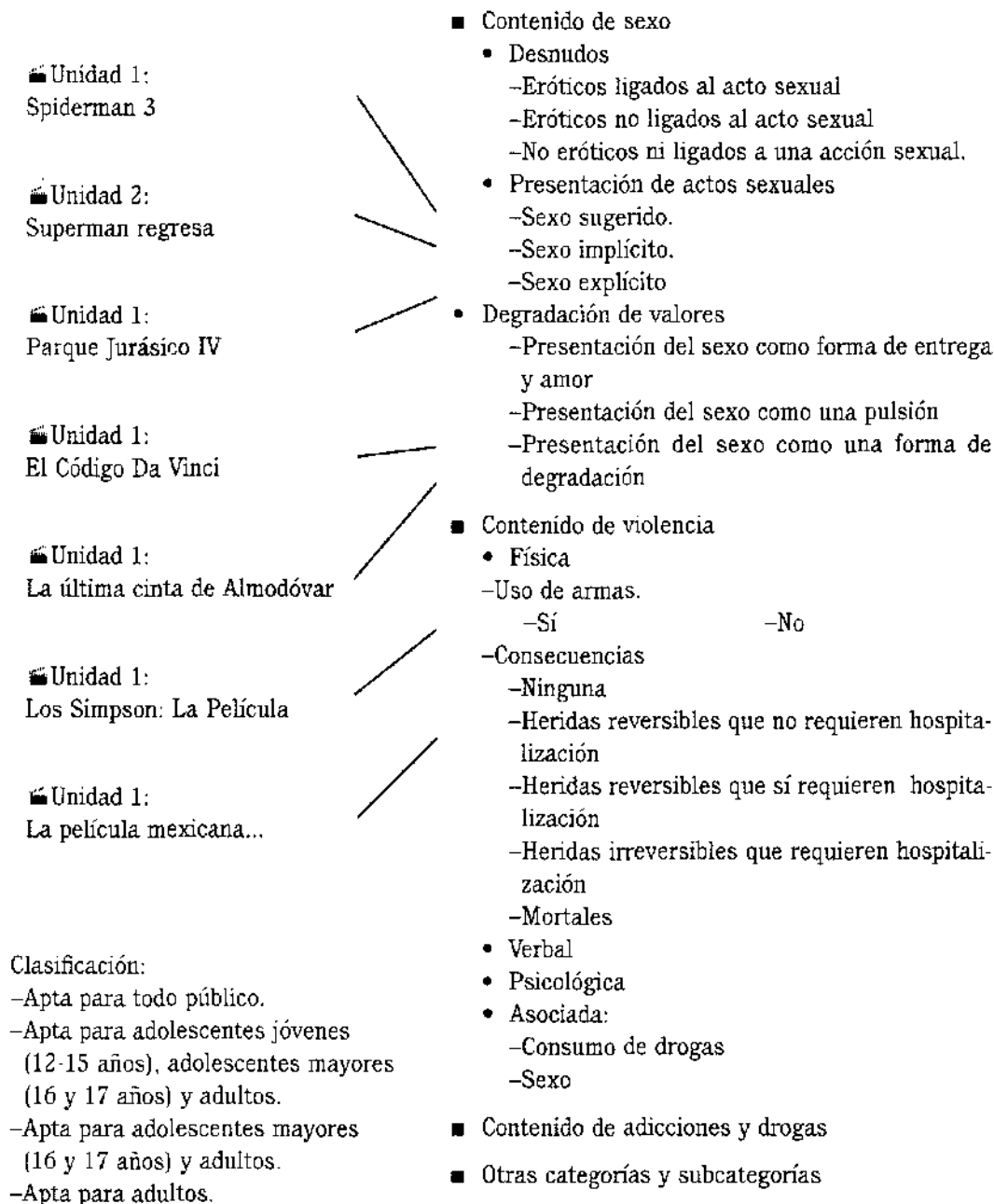
De esta forma, cada película se clasificaría en las categorías y subcategorías. Podríamos señalar que la presencia de ciertas subcategorías secundarias otorga un estatus u otro a la cinta en cuestión, por ejemplo:

Películas aptas para todo público

- Sólo se permiten desnudos no eróticos; si alguna escena presenta desnudos eróticos (ya sea que estén o no ligados al acto sexual), por lo que el filme se clasificará, según sea el caso, como "apto para adolescentes jóvenes (12-15 años)", o "apto para adolescentes mayores (16 y 17 años) y adultos".
- Únicamente se permite sexo sugerido; si alguna escena presenta sexo implícito o sexo explícito, el filme será clasificado según sea el caso.
- Imposible vincular el sexo con valores, mucho menos presentarlo como una pulsión instintiva y menos aún como una forma de degradación humana o humillación (subcategorías secundarias).

- La violencia únicamente puede presentarse en un contexto fantástico caricaturizado.
- Ningún contenido sobre adicciones y/o drogas está permitido.

Así, con cada grupo de edad: apta para adolescentes jóvenes (12-15 años); apta para adolescentes mayores (16 y 17 años) y adultos, y apta solamente para adultos o mayores de edad. Este proceso de categorización podría representarse gráficamente de la siguiente forma:



Para complementar el ejercicio le sugerimos que, junto con otros compañeros, lo termine y genere un sistema para clasificar películas (simplemente mostramos algunas categorías por las cuales puede comenzar; desde luego, podría no estar de acuerdo con ellas o con las subcategorías mencionadas, o con los criterios para otorgar un estatus u otro a la cinta). Una vez que desarrolle el sistema, vaya al cine, analice algunos filmes de estreno y otórgueles su propia clasificación de acuerdo con sus criterios.

El **análisis de contenido** transforma sistemáticamente material verbal, textual o visual en datos cuantitativos, mediante un conjunto de códigos y reglas para aplicarlo (Stark y Roberts, 2005).

¿Qué es y para qué sirve la observación cuantitativa?

Esta técnica de recolección de datos consiste en el registro sistemático, válido y confiable de comportamientos o conductas que se manifiestan. Puede utilizarse como instrumento de medición en muy diversas circunstancias. Haynes (1978) menciona que es el método más utilizado por quienes se orientan conductualmente. Puede servir para determinar la aceptación de un grupo con respecto a su profesor, así como para analizar conflictos familiares, eventos masivos (como la violencia en los estadios de fútbol), el rechazo de un producto en un supermercado, el comportamiento de personas con capacidades mentales distintas, etcétera.

La **observación cuantitativa**, como método para recolectar datos, es similar al análisis de contenido. De hecho, es una forma de observación del contenido de comunicaciones verbales y no verbales. De este modo, en este apartado algunos conceptos sólo se mencionarán pero no se profundizará en ellos, pues se tratan en el apartado sobre análisis del contenido.

Un aspecto importante a destacar es que en la observación el instrumento de recolección de los datos es una o varias personas, porque los equipos (de filmación, para captar movimientos, etc.) solamente registran lo que los sentidos humanos van a percibir (Anastas, 2005).

En la observación cuantitativa, a diferencia de la cualitativa, las variables a observar son especificadas y definidas antes de comenzar la recolección de los datos. Se enfoca en información que pueda ser evaluada por medio de los sentidos (datos visuales, auditivos, producto del tacto y el olfato). Asimismo, los observadores registran lo que perciben siguiendo reglas que se aplican invariablemente y deben minimizar su efecto sobre los registros e interacciones con los participantes observados (Anastas, 2005).

Observación cuantitativa. Registro sistemático, válido y confiable de comportamientos o conductas que se manifiestan en una circunstancia determinada. Se enfoca en información sobre las conductas más que en las percepciones.

Pasos para construir un sistema de observación

Los pasos para construir un sistema de observación son:

1. *Definir con precisión el universo de aspectos, eventos o conductas a observar.* ¿Qué y quién o quiénes se van a observar y por cuánto tiempo? Por ejemplo, si nuestro interés es conocer los recursos con que cuentan las escuelas de un distrito escolar, debemos definir con precisión

lo que concebimos como “recurso escolar”. Un universo podría ser el comportamiento verbal y no verbal de un grupo de alumnos durante un semestre en una materia. Otro universo lo constituirían las conductas de un grupo de trabajadores durante sus sesiones en círculos de calidad o equipos para la calidad, en el periodo de un año. O bien, las conductas agresivas de un grupo de esquizofrénicos en sesiones terapéuticas en cierto periodo. El universo obviamente depende de las variables que conformen el planteamiento del problema.

3. *Especificar las circunstancias de la observación.* ¿Dónde y cuándo se conducirán las observaciones? (Lugar y ubicación temporal). Por ejemplo, el aula en la clase de Economía durante este semestre desde el inicio y hasta que concluya el ciclo, antes del periodo de exámenes; todas las sesiones de los grupos la calidad del departamento de producción, en el lugar donde se reúnen normalmente; las próximas 10 sesiones terapéuticas de ciertos pacientes en una cámara de Gesell.²³ Tales circunstancias dependen del planteamiento del problema y la situación, los lugares pueden ser muy variados: una sala de emergencias, un consultorio médico, una tienda, etc. Resulta obvio que vamos a observar en lugares adecuados y evitar los que no son apropiados (por ejemplo, intentar registrar conductas criminales en un desfile familiar, no sería lo común, lo lógico). A veces se arregla o adapta un ambiente, de tal modo que sea lo suficientemente estable y se aproveche para realizar comparaciones entre conductas por medio de observaciones repetidas (tanto en el tiempo como entre individuos). Además, es indispensable definir si la observación se realizará “en vivo” (mientras ocurren las conductas, aspectos o eventos), o posteriormente (se filmarán los comportamientos y hechos que después se analizarán). Especificar el tipo de equipo necesario para efectuar las tareas de observación (si estamos observando células cancerosas, requeriremos de un potente microscopio; si se trata de planetas, un telescopio de muy largo alcance; el policía que observa a un grupo de pandilleros requerirá de binoculares y un micrófono especial que pueda registrar las conversaciones; con tiburones podría necesitarse un equipo que nos logre aislar de los escualos; etcétera).

Un caso maravilloso de observación por equipo altamente especializado lo ofrece el proyecto “Impacto Profundo” de la Administración Nacional Aeronáutica y Espacial (NASA, por sus siglas en inglés). La misión envió una nave espacial para colisionar al cometa Tempel 1 (lo cual implicó una investigación sumamente avanzada, el diseño de tecnología nunca antes vista, complejos cálculos matemáticos y proyecciones). Entre algunos de los objetivos del proyecto, podemos citar los siguientes: analizar la composición de los cometas y meteoros similares e indagar más sobre los orígenes del universo. La colisión calculada se produjo el 4 de julio del 2005 a 23 000 mph (NASA, 2005). La observación se realizó mediante cámaras fotográficas y de video que se ubicaron en la propia nave Deep Impact (nave de colisión) y en la nave Flyby, así como mediante algunos de los telescopios más desarrollados en el mundo. Un proyecto ideal para la observación del espacio. Además de que ahora los seres humanos sabemos que somos capaces de impactar cualquier meteorito o cometa que amenace nuestro planeta. El conocer la composición de estos astros, permitirá diseñar el proyectil adecuado.

²³ Es un espacio donde los participantes son colocados para observación, cuenta con dos secciones separadas por un cristal de una sola vista o unidireccional (que permite ver solamente de una sección a la otra, pero no de ésta hacia la primera). Una sección es regularmente una sala o mesa de trabajo (desde donde el cristal parece ser un espejo o un vidrio opaco) y la otra, la zona de observación (integrada por butacas, sillas o un aula escolar), en la cual se ve y se escucha todo lo que ocurre en la sala o mesa de trabajo (incluso puede filmarse la acción).

3. *Incluir todo el universo o extraer una muestra representativa de aspectos, eventos o conductas a observar.* Un repertorio que sea suficiente para observar conductas, eventos o cuestiones.
4. *Establecer y definir las unidades de observación.* Las unidades deben ser captadas por los sentidos. Por ejemplo, observar cada vez que se presenta una conducta agresiva de un determinado paciente; se analizará cada minuto si el alumno está o no atento a la clase de Economía (la materia se imparte dos veces a la semana, dos horas cada clase, de 7:00 a 9:00 horas), etc. El concepto de unidad de análisis es el mismo que en la recopilación y el análisis de contenido, sólo que en la observación cuantitativa se trata de conductas, eventos o aspectos.

Tabla 9.14 Algunos ejemplos de unidades de análisis o registro

Estudio	Unidades de observación/medición
Evaluar si un programa de televisión con contenido violento genera conductas agresivas en los niños.	Agresiones físicas de los niños (cada vez que se manifiesta una agresión como golpear a otro niño, empujarlo intencionalmente, etcétera).
Determinar la conducta de compra en una tienda de ropa por parte de cierto tipo de clientes.	Episodio de compra: Desde que ingresa a la tienda y hasta que sale de ésta. ¿Qué prendas o accesorios compró?, ¿de qué colores, texturas, etcétera?
Analizar la atención que se presta al tablero de avisos de la compañía.	Episodio de interacción: Cada vez que un empleado pasa junto al tablero de avisos, se detiene o no a verlo, cuánto tiempo pasa mirándolo y qué información capta más su atención.
Evaluar el efecto del consumo de alcohol en el comportamiento agresivo de los jóvenes.	Estancia en un bar o cervecería: Número de copas o tragos que consumen y número de comportamientos agresivos y sus características (registro progresivo de todo el tiempo de permanencia en el lugar).
Inferir actitudes de fervor religioso en la iglesia.	Cada periodo de tres minutos: registrar conductas que reflejen el fervor religioso (persignarse, arrodillarse, etcétera.)

5. *Establecer y definir las categorías y subcategorías de observación.* Estas categorías son similares a las definidas para el análisis de contenido. Y la observación también consiste en asignar unidades a categorías y subcategorías de observación. En éste, como en los demás métodos de medición, se asume que las variables (representadas por las categorías y subcategorías) han sido definidas conceptual y operacionalmente de forma adecuada, para que el observador realice correctamente las tareas de discriminación, que es su función central.

Tabla 9.15 Ejemplos de definiciones de categorías

Variable	Categorías	Definición
Conducta agresiva hacia los demás	0 a k conductas agresivas	Comportamientos que se manifiesten con la intención aparente de injuriar o dañar a otra persona, ya sea mediante insultos verbales o signos no verbales y acciones físicas.

(continúa)

Tabla 9.13 Ejemplos de definiciones de categorías (continuación)

Variable	Categorías	Definición
Conducta de compra (ropa)	0 a k artículos adquiridos	Compra de artículos a crédito o al contado.
Amabilidad en el servicio cuando el cliente tiene alguna necesidad	• Amable • Neutral • Grosero	Amable: Atender al cliente sonriéndole y haciendo contacto visual con él. Preguntarle por sus necesidades, resolver sus inquietudes, no interrumpirlo cuando esté hablando, dar solución a las necesidades pertinentes, mantenerse con él hasta que la solución solicitada se implemente, preguntar si requiere algo más e indagarle si está satisfecho. Neutral: No sonreír al cliente, hacer contacto visual solamente una vez o no mantener tal contacto. No demostrarle que se le está escuchando, no permanecer con él hasta que la solución se haya implementado y no verificar su satisfacción con dicha solución. Grosero: No sonreír al cliente, mostrarse molesto por sus peticiones, no establecer contacto visual, no preguntarle por sus necesidades, no resolver sus inquietudes, interrumpirlo cuando esté hablando, no solucionar las necesidades pertinentes, no mantenerse con él hasta que la solución solicitada se haya implementado, no preguntar si requiere algo más ni indagarle si está satisfecho.
Oportunidad en la atención al cliente (en una sucursal bancaria, en una taquilla en el cine, puesto de venta de boletos en un estadio de fútbol)	0 a k minutos y segundos	Tiempo que el cliente espera en fila (cola).

De unidades y categorías

En el caso del estudio citado (al hablar de la manipulación de variables independientes en experimentos) para probar la *hipótesis*: A mayor grado de información sobre la deficiencia mental que el sujeto normal maneje, mostrará menor evitación en la interacción con el deficiente mental (Naves y Poplawsky, 1984). Las unidades de análisis eran lapsos de 10 segun-

dos. La interacción entre la persona normal y el actor que hacía el papel de "deficiente mental" (individuo con capacidad mental distinta) duraba tres minutos. La variable dependiente fue "evitación de la interacción" y las categorías fueron cuatro (Naves y Poplawsky, 1984, pp. 107-109):

1. Distancia física: Se refiere a si el sujeto experimental aumenta o disminuye su distancia hacia el interlocutor a partir de la distancia que inicialmente ocuparía; esta distancia inicial estuvo delimitada por los asientos que el actor y el sujeto debían ocupar y, según la teoría, es la distancia en la que dos extraños en una situación de comunicación pueden interactuar de manera cómoda. Las dimensiones o subcategorías que este indicador adquiere son: el acercarse (afiliación) con valor de "2", el mantenerse estático, con valor de "1" o el alejarse (evitación) del actor (deficiente mental), con valor de "0", mediante inclinaciones corporales, o bien, al modificar por completo su distancia.
2. Movimientos corporales que denotan tensión: Esta categoría se orienta a captar los movimientos que el sujeto realiza como indicador de tensión (evitación) con valor de "0", ningún movimiento con valor de "1", que experimenta relajación (afiliación), con valor de "2". En este indicador específicamente se analizan si los movimientos de pies y piernas están a un ritmo acelerado, si hay ademanes con brazos y manos (como rascarse, sentir picor, etcétera) y la postura en general del sujeto.
3. Conducta visual del sujeto: Que según se estipula en esa investigación adquiere dos dimensiones o subcategorías:
 - a) dirigida hacia el interlocutor (afiliación) con valor de "1".
 - b) dirigida hacia cualquier otra parte (evitación) con valor de "0".
4. Conducta verbal: Tal indicador está compuesto por el contenido verbal del discurso del sujeto hacia el deficiente y se orienta primordialmente a la forma del discurso; incluye dos modalidades:
 - a) frases u oraciones completas (afiliación) con valor de "1".
 - b) frases dicotómicas y silencios (evitación) con valor de "0".

La modalidad de frases dicotómicas incluye respuestas monosilábicas (sí, no), murmullos, sonidos guturales; los silencios que se interpretan como respuestas dicotomas (respuesta de evitación) son los silencios no naturales en el discurso, cuando expresamente el sujeto se queda en silencio. La conducta verbal se mide a través del diálogo que sostenga el sujeto con el deficiente mental; es decir, en respuesta al guión que el actor interpreta ante cada sujeto (que es idéntico para todos) y en las intervenciones que el mismo sujeto realice.

En este caso, además de medir la conducta verbal en cuanto a su forma, convendría medirla también en cuanto a su contenido; es decir, determinar si las frases emitidas por el participante en respuesta a lo expresado por el deficiente revelan un contenido positivo o negativo.

Los periodos de observación, las unidades y categorías varían de acuerdo al planteamiento del problema e hipótesis.

Las subcategorías pueden ser escalas de actitudes

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

Figura 9.25

Categoría	Subcategorías
-Atención del alumno	Elevada (3) Media (2) Baja (1) Nula (0)
-Disciplina del alumno	Completa (3) Aceptable (2) Media (1) Indisciplina (0)
O bien agresividad del niño:	Muy alta (5) Alta (4) Regular (3) Baja (2) Inexistente (1)
O, al hablar del desempeño laboral observado en coordinadores de escuelas o asociaciones:	
Productividad	(5) (4) (3) (2) (1) (0) Improductividad
Actitud de servicio	(5) (4) (3) (2) (1) (0) Actitud de no servicio
Cumplimiento-incumplimiento	
Etcétera	

6. *Diseñar la manera de registrar las observaciones* (mediante un formulario, hoja de codificación o guía de observación). Es un instrumento para recolectar los datos, prediseñado sobre la base de la revisión de la literatura o construido especialmente para el estudio. Su formato es similar al que se presentó en el apartado sobre análisis de contenido. Anastas (2005) comenta que en ocasiones se utilizan sistemas de observación probados debido a: *a)* los costos (desarrollar uno requiere tiempo, varias pruebas y recursos), *b)* existen diversos instrumentos válidos y confiables para un amplio repertorio de conductas, *c)* se facilita la comparación con los estudios que utilizaron el mismo sistema de observación (se comparte un marco interpretativo). En el caso de Naves y Poplawsky (1984), la hoja de codificación es la que se muestra en la figura 9.25. asimismo, se agrega otro ejemplo en la figura 9.26.

Hora _____ Nombre del sujeto _____ Condición experimental _____

Codificador _____ Fecha _____

Categoría	Subcategoría	0'	10'	20'	30'	40'	50'	1'	10'	20'	30'	40'	50'	2'	10'	20'	30'	40'	50'	3'
Distancia física	Alejamiento (0)																			
	Acercamiento (2)																			
	Estático (1)																			
Movimientos corporales	Tensión (0)																			
	Relajación (2)																			
	Ninguno (1)																			
Conducta visual	Al sujeto (1)																			
	A otra parte (0)																			
Conducta verbal	F. completas (1)																			
	F. dicótomas (0)																			
	Silencios (0)																			

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

Figura 9.25 Primer ejemplo de una hoja de registro o codificación.

Los números (columnas) representan episodios de conducta inapropiada en la clase (cada vez que ocurre). Se coloca en la fila "CI" (conducta inapropiada) la letra inicial de la conducta (ejemplo: O, ofensa verbal; H, hablar en clase...), 1 si son otras. Asimismo, en la fila "N" se indica el número de estudiantes involucrados en el episodio.

El ejemplo de la figura 9.26 es un formato para registrar conductas inapropiadas en la clase y analizar cuáles son las que se presentan con mayor frecuencia y el número de estudiantes involucrados (Creswell, 2005).

La hoja de codificación, formato o guía debe acompañarse de las instrucciones para su llenado o registros y la mención de las unidades de observación, así como el tiempo que durará ésta. Debe evitarse un instrumento que sea tan complejo que sea difícil de codificar y utilizarse en la realidad.

7. *Determinar el papel de los observadores.* ¿Qué tan cercanos deben estar el observador y los observados? A este respecto mucho se ha escrito y los roles varían desde una distancia prudentemente lejana, hasta una cercanía completa.

CONDUCTAS INAPROPIADAS EN EL SALÓN DE CLASE O AULA

Fecha: _____

Hora: _____

Clase: _____

Profesor: _____

Número de estudiantes: _____

Grado: _____

Conductas inapropiadas (CI):

Agresión física

Ofensa verbal

Insulto no verbal (señas)

Hablar en clase

Ruido provocado

Desobediencia

No enfocado en la clase (distráido)

1-Otras

Número de estudiantes involucrados (N):

1-Individuo

2-Diada

3-Grupo pequeño (3 a 5)

4-Grupo grande (6 a 15)

5-Toda la clase

Estudiantes

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CI															
N															
No.	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
CI															
N															
No.	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
CI															
N															
No.	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
CI															
N															

Figura 9.26 Segundo ejemplo de una hoja de registro o codificación.

Anastas (2005) plantea dos roles fundamentales de la relación observador-observados: *a)* participante (donde hay relación entre el observador y las personas observadas) y *b)* no participante (la relación es mínima —simplemente por el hecho de que los individuos saben que están siendo observados— o puede ser nula si los individuos no saben que están siendo observados).

La ética de la investigación marca que siempre que va a observarse a un ser humano, se debe obtener su consentimiento; asimismo, no se deben invadir sus esferas privadas (baños, recámaras, etc.). Anastas (2005) está de acuerdo con esta consideración, aunque marca una excepción: cuando se observan conductas inocuas que pudieran ocurrir de manera natural en situaciones públicas y cuando las personas observadas no son identificadas o afectadas por los observadores. En última instancia, si el investigador busca evitar la “reactividad” en el proceso (que la presen-

cia de los observadores provoque reacciones poco o nada naturales en los participantes), una vez que se efectúan las observaciones, debe solicitársele a cada sujeto su autorización, mantener su nombre en el anonimato, nunca mostrar las filmaciones o registros y respetar su negativa, si es el caso.

Jorgensen (1989) sugiere tres actuaciones: a) *participante activo*, el observador interactúa plenamente con los observados, tiene un papel de miembro, es uno más de ellos (por ejemplo, si se trata de observar el compromiso de trabajadores en los grupos de calidad, el observador es un miembro del grupo con funciones específicas). b) *participante privilegiado*, no es un miembro del grupo, pero tiene acceso a información amplia y privada, los observados lo conocen y le tienen confianza y c) *observador limitado*, no tiene más que el papel de observador externo. Rogers y Bouey (2005) los clasifican en el siguiente continuo:

Participante completo: Total interacción, su rol primario es participar y generar la interacción.	Participante observador: Su rol primario es participar, pero casi a la par observa.	Observador participante: Su rol primario es observar, aunque tiene ciertas participaciones.	Observador completo: Nula participación, se limita a observar.
---	---	---	--

Patton (1980) sugiere que el observador debe ubicarse en varios continuos para establecer su papel (que es una ubicación aproximada y no exacta):

Papel como evaluador-observador		
Observador participante (participa activamente en tareas o situaciones)	Observador que participa sólo parcialmente	Observador no participante (externo)
Papel ante los demás (observados, miembros del ambiente, comunidad o contexto)		
Abierto (los demás conocen su papel y saben que están siendo observados)	Intermedio (algunos conocen su papel, y saben que están siendo observados, y otros no)	Encubierto (nadie conoce su papel ni que está siendo observado)
Duración de la observación (estancia del observador)		
Breve (una hora, un día)		Larga (meses, años)
Enfoque de la observación		
Enfocado (un elemento, un significado, una interacción en especial)		General (holístico: todos los elementos o las unidades)

Sin que de ninguna manera sea una norma, cuando la observación cuantitativa trabaja con personas, regularmente es no participante, encubierta o intermedia, de corta estancia y enfocada. Mientras que la observación cualitativa es participante, abierta, prolongada y general.

- 8. *Seleccionar a los observadores-codificadores.* Puesto que son las personas que habrán de codificar la conducta y deben conocer las variables, categorías y subcategorías. De acuerdo con Anastas (2005), las habilidades más importantes de un buen observador son:
 - Selectividad para registrar lo que observa.
 - Uso de todos los sentidos, particularmente ante lo que no se esperaba encontrar.
 - Capacidad para seguir las reglas de registro que involucran el procesamiento de información.
- 9. *Proporcionar entrenamiento a observadores-codificadores* (en las variables, categorías, subcategorías, unidades de análisis y el procedimiento para codificar, así como sobre las diferentes maneras como puede manifestarse una categoría o subcategoría de conducta). Debido a que, en algún grado, la observación involucra hacer juicios, el entrenamiento debe apoyar la estandarización de procedimientos y, por ende, a la objetividad.
- 10. *Efectuar la prueba piloto y calcular la confiabilidad de los observadores* (intraobservador e interobservadores). En la prueba piloto se evalúa todo lo relativo al proceso de observación. Desde la iluminación en el ambiente (ejemplo, cámara de Gesell), el sonido y los equipos de filmación; hasta el lugar donde se coloquen los observadores, las instrucciones y la guía de observación o registro.

Gracias a la prueba piloto nos podemos dar cuenta de problemas potenciales y corregirlos (por ejemplo, si vamos a observar conductas agresivas en niños y nada más hemos trabajado con adultos, nos percataremos de que la agresión no se manifiesta igual en adultos que en niños —o adolescentes—. De este modo, durante la prueba piloto podemos registrar comportamientos agresivos no contemplados, que se agregarán como subcategorías en el instrumento de observación definitivo). Esto se hace con una muestra del repertorio de conductas o actos a considerar. Con respecto a la confiabilidad, recordemos que debe demostrarse y no asumirse y se determina para evaluar si los observadores están registrando coherentemente (confiabilidad individual) y si están registrando entre sí de manera similar los mismos eventos o conductas (confiabilidad interobservadores). Los procedimientos y las fórmulas pueden ser las mismas que las vistas en el apartado sobre el análisis de contenido; lo único que cambia es la palabra “codificadores”, “codificación”, “codificadas”, por “observadores”, “observación”, “observadas”.

Por ejemplo:

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

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

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

$$Ao = \frac{Ia}{Ia + Id}$$

Donde Ia es el número total de acuerdos entre observadores, e Id es el número total de desacuerdos entre observadores. Un "acuerdo" se define como la codificación de una unidad de análisis en una misma categoría por distintos observadores. Se interpreta como cualquier *coeficiente de confiabilidad* (0 a 1).

Es muy importante al establecer la confiabilidad, evitar que los observadores "copien" o vean el trabajo de sus compañeros, ya que esto puede ocurrir.

Anastas (2005) considera como una confiabilidad interobservadores mínima aceptable un 0.85 u 85%. Si se elige un instrumento ya desarrollado, al igual que en otros métodos de medición, debe demostrarse que en los estudios donde se aplicó resultó válido y confiable, y adaptarse a las condiciones de nuestra investigación.

Para establecer la validez de contenido, se analiza el dominio de nuestras categorías y subcategorías contra el universo de posibles conductas que pueden manifestarse. La validez de criterio muchas veces se establece usando otra técnica (por ejemplo, la hostilidad, que se infiere de observar actos agresivos, puede validarse mediante una escala que mida tal variable).

- 1.1 *Llevar a cabo la codificación por observación.*
- 1.2 *Vaciar los datos de las hojas de codificación y obtener totales para cada categoría.*
- 1.3 *Realizar los análisis apropiados.*

Ventajas de la observación

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

1. *Son técnicas de medición no obstrusivas.* En el sentido que el instrumento de medición no "estimula" el comportamiento de los participantes (las escalas de actitud y los cuestionarios pretenden "estimular" una respuesta a cada ítem). Los métodos no obstrusivos simplemente registran algo que fue estimulado por otros factores ajenos al instrumento de medición.
2. *Aceptan material no estructurado.*
3. *Pueden trabajar con grandes volúmenes de datos (material).*

Otras formas cuantitativas de recolección de los datos

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

En la investigación disponemos de otros métodos para recolectar los datos, tan útiles y fructíferos como los cuestionarios, las escalas de actitudes, el análisis de contenido y la observación, entre éstos se encuentran:

1. *Pruebas estandarizadas.* 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, calidad de vida, lealtad a una marca de bebidas refrescantes (refrescos o sodas), diferenciación de marcas (en productos concretos), etcétera.

Asimismo, hay un tipo de pruebas estandarizadas que miden 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 sujetos y éstos relatan sus asociaciones e interpretaciones en relación con las manchas).

Estas pruebas se consideran en el capítulo 7 "Recolección de los datos cuantitativos", segunda parte, incluido en el CD anexo.

2. *Datos secundarios* (recolectados por otros investigadores). Implica la revisión de documentos, registros públicos y archivos. 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 acudiríamos a los archivos de los hospitales y las diferentes procuradurías o cuerpos policíacos. Otro ejemplo sería consultar los archivos de una universidad y tomar los datos de los alumnos con respecto a la inteligencia, la personalidad u otras variables que nos interesen.

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 la 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.

La recolección de datos secundarios (incluyendo desde luego, estadísticas y metaanálisis), se presenta en el capítulo 7 "Recolección de los datos cuantitativos", segunda parte, incluido en el CD anexo.

3. *Instrumentos mecánicos o electrónicos.* Sistemas de medición por aparatos como por ejemplo: el detector o polígrafo de mentiras que considera la respuesta galvánica de la piel, la pistola láser que mide la velocidad a la que circula un automóvil desde un punto externo al vehículo, instrumentos que captan la actividad cerebral, el escáner que mide con exactitud el cuerpo de un ser humano y ubica la talla ideal para confeccionar toda su ropa o vestuario (muy socorrido por la milicia norteamericana para diseñar los uniformes de los soldados), la medición eléctrica de distancias, etc. Tales instrumentos se comentarán brevemente en el capítulo 7 "Recolección de los datos cuantitativos", segunda parte, incluido en el CD anexo.
4. *Instrumentos específicos propios de cada disciplina.* 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.). Para el análisis de grupos se utilizan los sistemas sociométricos y el análisis de redes, entre otros, algunos de estos métodos se comentan en el capítulo 7 del CD anexo y en el sitio de Internet de este libro.

¿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 y las categorías y subcategorías de contenido u observación requieren codificarse con símbolos o números; y deben codificarse, 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”). 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 antes de que éste se aplique) y que las preguntas abiertas no suelen estar precodificadas. Pero, en cualquier caso, una vez que se tienen las respuestas, éstas deberán codificarse.

La codificación de las respuestas implica cuatro pasos:

1. Establecer los códigos de las categorías o alternativas de respuesta de los ítems o preguntas y las categorías y subcategorías de contenido u observación no precodificadas.
2. Elaborar el libro de códigos.
3. Efectuar físicamente la codificación.
4. Grabar y guardar los datos y su codificación en un archivo permanente.

Veamos cada paso con algunos ejemplos.

1. Establecer códigos

Si todas las categorías y subcategorías fueron precodificadas y no se tienen preguntas abiertas, el primer paso no es necesario. Éste ya se efectuó.

Si las categorías no fueron precodificadas y se tienen preguntas abiertas, deben asignarse los códigos o la codificación a todas las categorías de ítems, de preguntas o de contenido u observación. Obviamente en este punto, cuando nos referimos al análisis de contenido y observación, hablamos también de las subcategorías. Por ejemplo:

Pregunta no precodificada:

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

☐ Sí ☐ No

Se codifica:

1 = Sí

0 = No

Frase no precodificada:

Creo que estoy recibiendo un salario justo por mi trabajo.

- | | |
|-------------------------------------|-------------------|
| () Totalmente de acuerdo | () De acuerdo |
| () Ni de acuerdo, ni en desacuerdo | () En desacuerdo |
| () Totalmente en desacuerdo | |

Se codifica:

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

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

2. Elaborar el libro o documento de códigos

Una vez que están codificadas todas las categorías del instrumento de medición, se procede a elaborar el “libro de códigos”.

El libro o documento de códigos describe la localización de las variables y los códigos asignados a los atributos que las componen (categorías o subcategorías) (Babbie, 2001). Este libro es la guía para: a) el proceso de codificación y b) para localizar variables e interpretar los datos durante el análisis (Babbie, 1995, 2001). El libro de códigos puede conducirnos a los valores de las categorías, de las variables, así como a sus significados; incluso, hace referencia a una matriz de datos. Veámoslo por partes.

Los elementos comunes de un libro de códigos son: variable, pregunta, ítem, tema, categoría general, categorías y subcategorías, códigos (números o símbolos utilizados para asignarse o designar a las categorías, columnas (número de la columna en la matriz de datos).

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 9.16.

Tabla 9.16 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 acuerdo	3	
		- En desacuerdo	2	
		- Muy en desacuerdo	1	
	Frase 2 (servicios)	- Muy en desacuerdo	5	2
		- De acuerdo	4	
		- Ni de acuerdo, ni en desacuerdo	3	
		- En desacuerdo	2	
		- Muy en desacuerdo	1	
	Frase 3 (deshonestidad)	- Muy de acuerdo	1	3
		- De acuerdo	2	
		- Ni de acuerdo, ni en desacuerdo	3	
		- En desacuerdo	4	
		- Muy en desacuerdo	5	

En el caso del estudio por observación de Naves y Poplawsky (1984) (figura 9.25), el libro de códigos sería el que se muestra en la tabla 9.17. Pedimos al lector lo revise detenidamente.

Es decir, el libro o documento de códigos es un manual para el investigador y los codificadores. En el que se encuentran las respuestas a los cuestionarios empleados, las escalas y pruebas administradas; los registros de las hojas de codificación de contenido u observación, o las mediciones de cualquier otro instrumento aplicado (los datos recolectados). Estos elementos son transferidos a una matriz, la cual es el conjunto de datos simbólicos o numéricos producto de la aplicación del instrumento. Esta matriz es lo que habrá de analizarse.

La matriz adquiere significado por el documento de códigos. El apartado columna dentro del libro de códigos tiene sentido en la matriz; veamos por qué. La matriz tiene renglones y columnas, los renglones representan casos o participantes en la investigación y las columnas son los ítems. Los casos (renglones o filas) van en orden progresivo descendente (1, 2, 3, ..., k), las columnas (ítems) van en orden progresivo de izquierda a derecha (1, 2, 3, ..., k). Cada ítem ocupa una columna. Los datos (respuestas o registros) se colocan en las celdas (intersección de un caso y un ítem, renglón y columna). Si en el libro de códigos se señala para un ítem o categoría

Tabla 9.17 Ejemplo de un documento de códigos en el caso del estudio de Naves y Poplawsky²⁴

Variable	Categorías	Subcategorías	Códigos	Columnas
– Tratamiento experimental	– Grupo cultural		1	1
	– Grupo sociopsicológico		2	
– Conducta	– Distancia física	– Alejamiento	0	2
		– Acercamiento	2	
		– Estático	1	
	– Movimientos corporales	– Tensión	0	3
		– Relajación	2	
		– Ninguno	1	
	– Conducta visual	– Al sujeto	1	4
		– A otra parte	0	
	– Conducta verbal	– Frases completas	1	5
		– Frases dicótomas o silencios	0	
– Codificador	– LRE	–	1	6

²⁴ Desde luego, Naves y Poplawsky (1984) para las categorías de “conducta de evitación” obtenían esta codificación cada 10 segundos (cada unidad de análisis), y sumaban el número de 2 (dos) y 1 (uno) y lo transformaban en porcentaje. Aquí suponemos que toda la interacción con el supuesto deficiente mental puede categorizarse y subcategorizarse. Recordemos que eran 30 participantes.

una cierta columna (por ejemplo, 10), significa que en la décima columna de la matriz se habrán de anotar las respuestas para dicho ítem. La matriz se esquematiza como se muestra en la figura 9.27.

Matriz de datos			
	Columna 1	Columna 2	Columna 3
Caso 1			Valores de categorías/ subcategorías
Caso 2			
Caso 3			
Caso 4			
Caso k			

Figura 9.27 Ejemplo de matriz de datos.

Los resultados del instrumento de medición se transfieren a la matriz, usando el libro o documento de códigos. El proceso se representa en la figura 9.28.

Sin el libro de códigos no puede llevarse a cabo la transferencia. Vamos a suponer que hubiéramos aplicado la escala de actitud con los tres ítems de la tabla 9.16 a cuatro personas, y así, obtuviéramos los resultados que se observan en la figura 9.29.

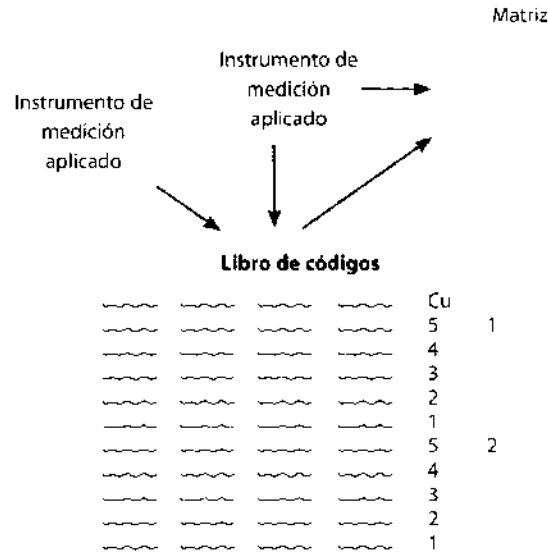


Figura 9.28 Proceso de transferencia de los datos a la matriz.

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

Figura 9.29 Ejemplo de aplicación de tres ítems a cuatro sujetos.

De acuerdo con el libro de códigos (tabla 9.16), tendríamos la matriz de la figura 9.30.

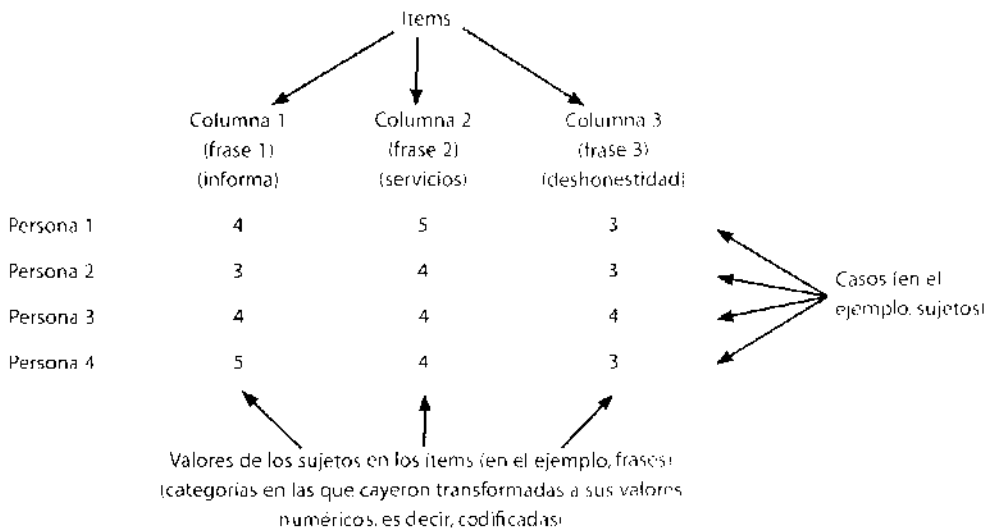


Figura 9.30 Ejemplo de matriz de datos para el libro de códigos de la tabla 9.16.

En el ejemplo de Naves y Poplawsky (tabla 9.17), la matriz sería como la que se muestra en la figura 9.31.

	Columna 1 Tratamiento Experimental	Columna 2 Distancia física	Columna 3 Movimientos corporales	Columna 4 Conducta visual	Columna 5 Conducta verbal	Columna 6 Codificador
S ₁	1	0	0	0	0	1
S ₂	1	0	0	0	0	1
S ₃	2	2	2	1	1	2
S ₄	1	1	0	0	0	2
S ₅	2	2	2	1	1	2
S ₆	1	0	1	1	0	3
S ₇	2	2	2	1	1	3
S ₈	2	2	2	1	1	3
S ₉	2	1	2	0	1	1
S ₁₀	1	2	1	0	0	2

Figura 9.31 Ejemplo hipotético de matriz de datos para el libro de códigos de la tabla 9.17 (Naves y Poplawsky).

El libro de códigos indica a los codificadores qué variable, ítem, categoría, subcategoría va en cada columna, y qué valores debe anotar en cada columna, así como el significado de cada valor numérico.

Con el libro de códigos sabemos que el sujeto 1 pertenece al "grupo cultural" ("1" en la primer columna), que tuvo una conducta de alejamiento en su distancia física ("0" en la segunda columna), que sus movimientos corporales fueron de tensión ("0" en la tercera columna), que su conducta visual fue a otra parte y no vio al sujeto ("0" en la cuarta columna), que dijo frases dicotómicas o silencios ("0" en la quinta columna) y que el codificador fue LRF ("1" en la sexta columna). Y así con cada sujeto.

Obsérvese que, sin el libro de códigos, no se puede codificar y una matriz de datos carece de significado. La siguiente matriz no nos dice nada sin el libro de códigos (ver tabla 9.18).

Tabla 9.18 Modelo de matriz de datos

5	0	0	2	1	1	0	1	0	1	5
5	0	1	2	2	3	2	3	0	0	4
2	1	0	0	0	0	0	0	0	0	0
1	1	1	1	1	4	0	2	0	0	3
3	0	0	1	2	5	0	2	0	0	3
3	1	0	1	2	5	0	1	0	1	3
4	0	1	0	2	4	0	1	0	1	4
5	1	1	2	2	1	0	0	0	1	4
2	0	1	2	2	2	0	2	1	0	5

¿Qué significa cada columna, cada dígito? Está en clave y sólo podemos tener acceso a ella mediante el libro de códigos.

Valores perdidos

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 (por ejemplo, no se pudo observar la conducta), se crean una o varias categorías de valores perdidos y se les asignan sus respectivos códigos.

EJEMPLO

Sí = 1		Sí = 1
No = 2		No = 2
No contestó = 3		
Contestó incorrectamente = 4	o	Valor perdido por diversas razones = 9

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.

Hasta el momento se han presentado, por razones didácticas, ejemplos resumidos de libros de códigos. Desde luego, un libro de códigos casi siempre tiene más variables o categorías y, en

consecuencia, más columnas (al igual que la matriz de datos). Hay matrices que pueden tener 500 o más columnas. Asimismo, debe recordarse que los renglones son casos (sujetos, escuelas, series de televisión, etcétera).

Actualmente el *libro de códigos* cumple una función más bien de comunicación entre los investigadores que participan en un estudio, para que todos sepan las claves de los códigos asignados a los datos; ya que en la mayoría de los programas de análisis estadístico (como el Paquete Estadístico para las Ciencias Sociales, SPSS) la matriz de datos va acompañada de la matriz de códigos. Veamos este caso con dicho programa, que es uno de los más usados en la investigación. Otros programas siguen un patrón similar.

SPSS tiene dos ventanas: *a)* vista de los datos y *b)* vista de las variables. La primera aparece si seleccionamos en la “vista de datos”, que es la pestaña (simulación de una carpeta o *fólder*) ubicada en la parte inferior de la pantalla hacia nuestro lado izquierdo. La segunda está hacia la derecha y es la pestaña que dice “vista de variables”.

La “vista de datos” es la matriz de datos (renglones = casos; columnas = ítems, categorías, subcategorías; celdas = datos). En esta vista se introducen los datos (valores).

La “vista de variables” es el libro o documento de códigos electrónico (tiene también forma de matriz). Pero en esta “vista”, los renglones o filas significan ítems o categorías y las columnas representan características de cada ítem o categoría. 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 investigación y una columna en la matriz o “vista de los datos”, y un renglón o fila en la “vista de variables”) y en otras ocasiones son un ítem, reactivo, categoría o subcategoría de una variable de investigación).

Conceptualmente tenemos: en la vista de variables se señala —ítem por ítem, variable por variable de la matriz— los siguientes elementos (que son columnas):

1. Nombre de cada ítem o variable de la matriz (cada vez que hablamos de ítem nos referiremos también a categorías de contenido u observación o a variables de investigación integradas por un único reactivo).
2. Tipo de variable (numérica, no numérica o cadena —símbolos, números que indican un nivel nominal—, cantidad, fecha, etc.). Incluso una cifra con decimales. Este tipo se vincula al nivel de medición.
3. Anchura (en dígitos). Esto depende de la comodidad de ancho con la cual deseemos trabajar y del ancho de las categorías (ejemplos: en un ítem de actitud la calificación ocupa un dígito —totalmente de acuerdo = 5, de acuerdo = 4, etc.—, ingresos puede ocupar varios dígitos de acuerdo al tipo de moneda y si no agrupamos y decidimos colocar la cantidad completa).
4. Decimales (si es pertinente).
5. Etiqueta (definición o párrafo que define a la variable en la matriz o ítem).
6. Valores. Los códigos de cada opción de respuesta o subcategoría. La codificación en sí. Incluye, desde luego, valor y su etiqueta. También de los valores perdidos.
7. Los 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 quinto decimal si es pertinente).

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: escala —intervalo o razón—, ordinal o nominal).

En la figura 9.32 mostramos un ejemplo de la vista de variables.

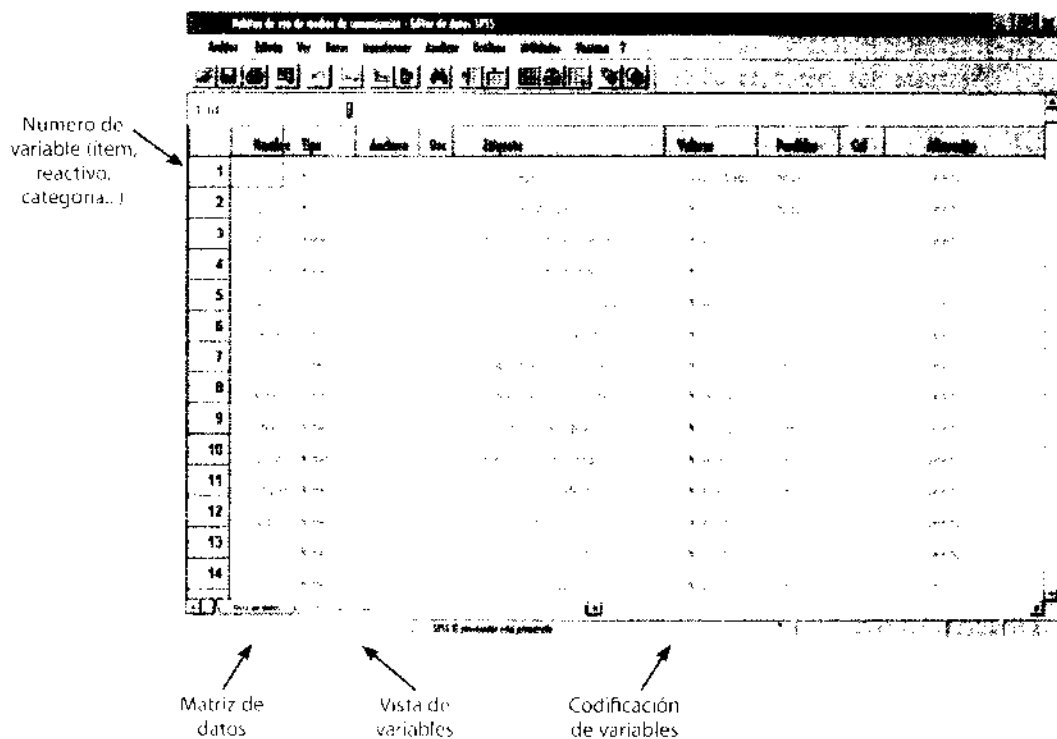


Figura 9.32 Vista de variables en SPSS.

3. Codificación física

El tercer paso del proceso es la codificación física de los datos, es decir, llenar la matriz de datos con valores (en SPSS, vista de los datos). Esta tarea la efectúan los codificadores, a quienes se les proporciona el libro de códigos. Así, cada codificador vacía las respuestas en la matriz de datos, de acuerdo con el libro de códigos. El vaciado de la matriz de datos puede hacerse en "hojas de tabulación", las cuales tienen columnas y renglones. En la figura 9.33 se muestra un ejemplo de una de estas hojas.

Fecha _____

Investigación _____

Codificación _____

Hoja No. _____

Diagrama de una hoja de tabulación:

Columnas (variables) →

Renglones (casos) →

Figura 9.33 Ejemplo de una hoja de tabulación.

Si no alcanza una hoja de tabulación, se utilizan las hojas necesarias para vaciar los datos de todos los casos. Por ejemplo, la primera hoja puede alcanzarnos para 24 casos; pero si tenemos 200 casos, habremos de usar nueve hojas. Cada hoja estará llena de dígitos. Desde luego, actualmente las hojas son matrices u hojas de cálculo (por ejemplo, en Excel).

Hoy día, en la mayoría de las investigaciones ya no se usan las hojas de codificación. Lo común es transferir los datos o valores directamente de los ejemplares del instrumento de medición a la matriz de datos creada como archivo en la computadora (vista de datos en SPSS, hoja de cálculo como Excel que después se copia en SPSS, matriz de entrada en Minitab y similares). Esta tarea de captura debe realizarse de manera muy cuidadosa, para evitar cometer errores y alterar los resultados.

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

4. Generación de archivo o archivos

En cualquier caso se crea la matriz de datos y se archiva, así se genera un archivo, el cual debe ser nombrado y contener los datos codificados en valores numéricos.

Es muy importante revisar la matriz o vista de los datos ("limpiarla de errores"). Verificar que no haya errores de codificación, valores que no correspondan (imaginemos que en un ítem de Likert que se codifique del cinco al uno aparezca un 7 o un 35), respuestas duplicadas, etcétera.

Se puede limpiar: *a)* revisándola físicamente, *b)* en SPSS con la función "ordenar casos" (en "Datos") y, de este modo, visualizar valores que no correspondan a cada variable de la matriz (o ítem), *c)* ejecutando el "análisis de frecuencias" en el menú "Análizar" y "estadísticos descriptivos", 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.

Hemos generado un archivo con los datos recolectados y codificados. El archivo puede almacenarse en el disco duro, un disquete o un CD (disco compacto). El proceso se muestra en la figura 9.34.

Así, los datos han sido capturados en un archivo permanente y están listos para analizarse mediante un programa de computadora. El proceso va desde la respuesta de los participantes hasta un archivo que contiene una matriz (que es una matriz de valores numéricos que significan respuestas). De esta forma, los valores de las variables se han codificado.

¿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 ayuda a establecerse la validez de criterio. No solamente se puede, sino que es conveniente, hasta donde lo permita el presupuesto para investigar.

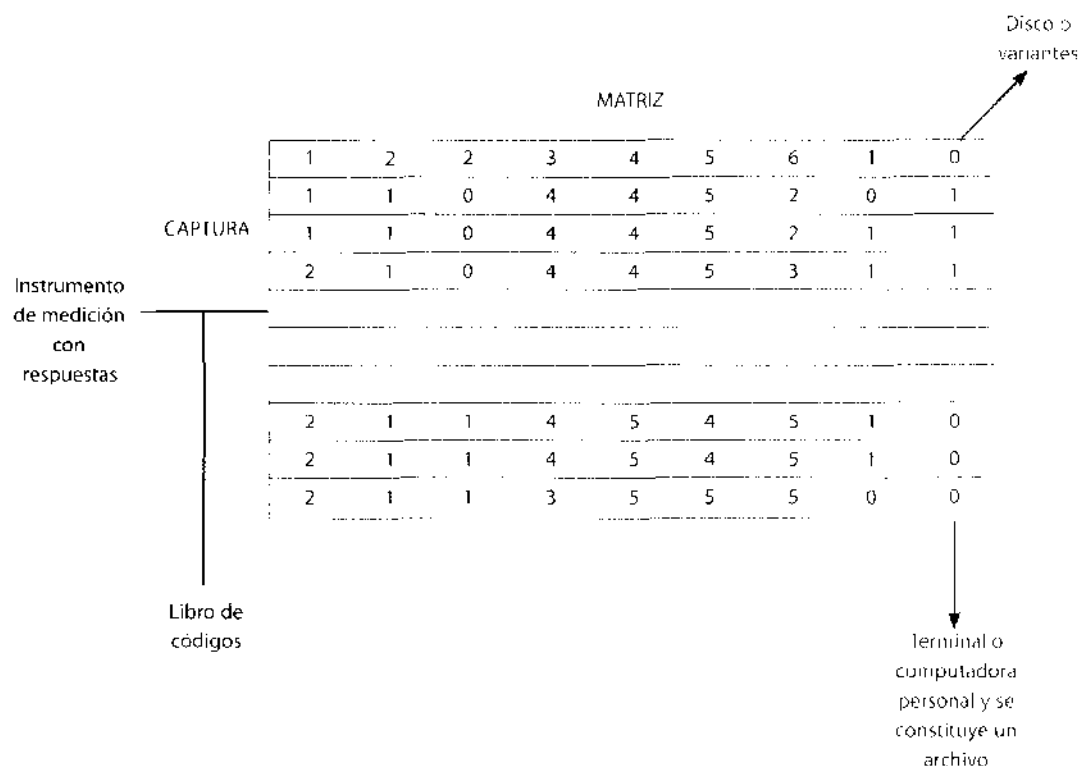


Figura 9.34 Proceso para crear un archivo de datos (la matriz de datos).

- 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, al mismo sujeto u objeto, 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 el sujeto participante 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 confiabilidad.
- Los coeficientes de confiabilidad 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 inicial).
 7. Construir 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 Guttman.
 2. Cuestionarios (autoadministrado, por entrevista personal, por entrevista telefónica y por correo).
 3. Recolección de contenidos para análisis cuantitativo.

4. Observación cuantitativa.
5. Pruebas estandarizadas (procedimiento estándar).
6. Archivos y otras formas de medición.

Algunos se incluyen en el capítulo 7 del CD anexo.

- Las respuestas a un instrumento de medición se codifican.
- La codificación implica:
 1. Codificar los ítems o equivalentes no precodificados.
 2. Elaborar el libro o documento de códigos.
 3. Efectuar físicamente la codificación.
 4. Grabar y guardar los datos en un archivo permanente.
- Para resumir algunos de los instrumentos tratados en el capítulo se agrega la tabla 9.19:

Tabla 9.19 Concentrado de instrumentos para la recolección de datos

Métodos	Propósito general básico	Ventajas	Retos
Cuestionarios/ Escala de actitudes/ Pruebas estandarizadas	–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).	–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.	–Regularmente no se obtiene retroalimentación detallada de parte de los sujetos. –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	–Recolectar información no obstrusiva respecto a conductas y procesos.	–Se puede adaptar a los eventos tal y como ocurren. –Se evalúan hechos, comportamientos y no mediciones indirectas.	–Dificultad para interpretar conductas. –Complejidad al categorizar las conductas observadas. –Puede ser obstrusiva y provocar sesgos si es “participante”. –Puede ser costosa.

(continúa)

Tabla 9.19 Concentrado de instrumentos para la recolección de datos (*continuación*)

Métodos	Propósito general básico	Ventajas	Retos
Análisis de contenido	-Recolectar información no obstrusiva respecto de mensajes.	-Se puede adaptar a los eventos tal como ocurren. -Se evalúan mediciones indirectas.	-Dificultad para interpretar mensajes. -Complejidad al categorizar los mensajes.

CONCEPTOS BÁSICOS

Análisis cuantitativo de contenido
 Archivo de datos
 Autoadministración
 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
 Entrevista
 Evidencia relacionada con el constructo

Evidencia relacionada con el contenido
 Evidencia relacionada con el criterio
 Hojas de codificación
 Hojas de tabulación
 Instrumento de medición
 Matriz de datos
 Medición
 Medida de estabilidad
 Método de formas alternas
 Método de mitades partidas
 Niveles de medición
 Observación cuantitativa
 Pruebas estandarizadas
 Pruebas proyectivas
 Recolección de datos
 Unidad de análisis
 Validez

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. Suponga que alguien intenta evaluar la actitud hacia el presidente de la República, y 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).

5. Construya un cuestionario para medir la variable que considere conveniente e incluya preguntas demográficas, por lo menos 10 preguntas más. Aplíquelo a 20 conocidos suyos; elabore el libro de códigos y la matriz de datos, vacíela en una hoja de tabulación elaborada por usted. Finalmente, obtenga de la hoja de tabulación, el significado de los dígitos de todas las columnas correspondientes a los cinco primeros casos.
6. ¿Cómo mediría la hostilidad mediante observación y cómo por medio de una escala de actitudes?
7. Genere un planteamiento del problema, donde utilice por lo menos dos tipos de instrumentos cuantitativos para recolectar datos.
8. ¿Cómo se podría aplicar el análisis cuantitativo del contenido para la evaluación de un programa educativo a nivel superior?
9. 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 la últimas elecciones municipales y líder histórico al que más admira. Que los participantes sean sus compañeros de clase (el ejercicio implica levantar datos y codificarlos, desde luego).
10. Elabore uno o varios instrumentos para el ejemplo de estudio que ha desarrollado hasta ahora en el proceso cuantitativo (incluya la codificación).

EJEMPLOS DESARROLLADOS²⁵

LA TELEVISIÓN Y EL NIÑO

Se aplicó un cuestionario en una muestra total de 2 112 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 televi-

sión, 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. El cuestionario se muestra en el capítulo 7, "Recolección de datos cuantitativos, segunda parte" del CD anexo.

EL CLIMA ORGANIZACIONAL

En una primera etapa, el cuestionario fue aplicado en: 1) una empresa transnacional de la industria de la transformación con

²⁵ Por cuestiones de espacio, se comentan brevemente.

cerca de 500 empleados (laboratorio químico-farmacéutico), 2) una empresa mexicana del mismo tipo de industria con más de 500 empleados, 3) una institución de educación media y superior con más de 200 personas empleadas y 4) una empresa comercial con poco más de 100 empleados.

Las dimensiones consideradas para todas las muestras fueron: moral, apoyo de la dirección, innovación, percepción de la empresa-identidad-identificación, comunicación, percepción del desempeño, motivación intrínseca, autonomía, satisfacción general, liderazgo, visión y recompensas o retribución. Recordemos que lo que se evalúa son las percepciones sobre tales atributos organizacionales.

En el caso de la institución educativa, se incluyeron también dichas variables, con excepción de recompensas, debido a que la alta dirección no consideró prudente evaluarla.

Se utilizaron dos tipos de reactivos: 1) Frase que busca la reacción del sujeto y que se refleja en una alternativa de respuesta (primera parte) y 2) pregunta directa, cuyas categorías ubican al sujeto en alguna posición (segunda parte). Se estableció un escalamiento de tipo Likert con cinco opciones de respuesta.

La primera versión del cuestionario incluyó 76 ítems en forma de frase y 24 preguntas. Ésta fue sometida a una prueba piloto de comprensión, lenguaje y confiabilidad inicial con 62 sujetos de una empresa mediana que fabrica dulces y otra dedicada a elaborar calzado. El coeficiente alfa-Cronbach obtenido fue de 0.91 y 0.92 ($p < 0.05$).

Después de la prueba piloto, se eliminaron tres reactivos tipo frase y una pregunta. Así, el instrumento sometido a validación constó de 73 frases y 23 preguntas, el cual se presenta en el capítulo 7 "Recolección de los datos

cuantitativos", segunda parte (incluido en el CD anexo). Cabe señalar que el instrumento fue "personalizado" con el nombre de la empresa o institución en la cual se aplicó.

La confiabilidad promedio fue de 0.95 (alfa-Cronbach). El proceso de validación se comentó a lo largo de este capítulo (Hernández Sampieri, 2005). La validación de constructo se efectuó en cada muestra por análisis de factores, el cual se muestra en el capítulo 8 del CD anexo: "Análisis estadístico-multivariado de los datos".

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 lo tanto, tal inconveniente es intrascendente en el caso de estudio. Otra

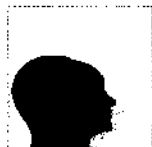
desventaja que se le atribuye, es que la escala no incluye elementos que evalúen la actitud de los menores ante los contactos positivos, para determinar si los PPASI (programas de prevención de abuso sexual infantil) generan un efecto nocivo de suspicacia indiscriminada ante cualquier contacto. Por lo cual, se decidió incluir un par de reactivos para evaluar esta posibilidad en el Role Play-México (RP-México). 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 aborda 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 RP-México 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 (RP-México 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 contactos inapropiados (seis máximo) y un punto por cada persona a quién 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 a un método de formas paralelas al administrarse el RP-México y el EPA. La correlación entre ambas pruebas alcanzó un buen nivel y fue signi-

ficativo ($r = 0.75, p < 0.01$), 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, RP-México y en todas las subescalas, los índices van de 0.59 a 0.78, todas con $p < 0.01$. El instrumento RP-México 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 DEBEN

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 instru-

mento. Si bien el ejemplo puede resultar un tanto simple y grosero, en el nivel de las investigaciones de pre-grado, 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 corre este riesgo, el investigador podría llevar sus conclusiones más allá de la realidad, cuando lo que se desea es reflejar la realidad.

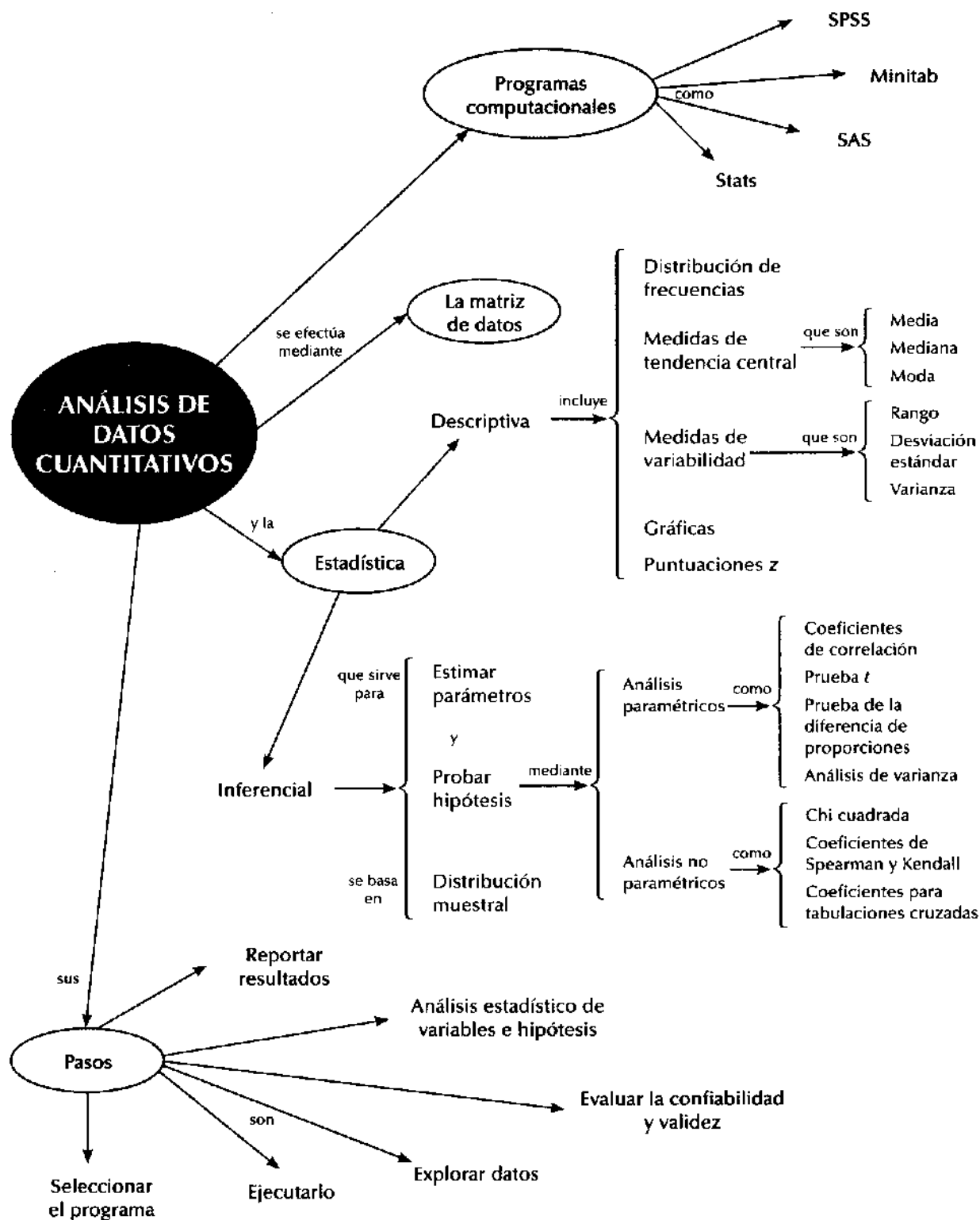
Un idea clave, para no tropezar con asuntos insalvables en este momento de la investigación o para no tomar decisiones que conduzcan al error, es hacer un buen proyecto de investigación. En la etapa de la planificación debe quedar claramente establecido y justificado qué

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.

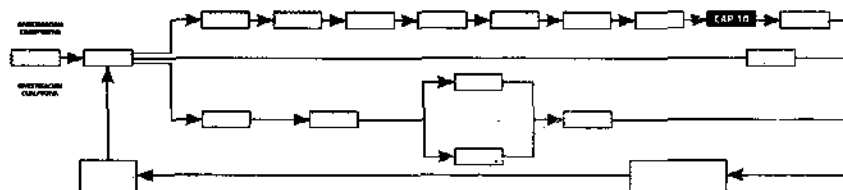
Edwin Salustio Salas Blas

*Universidad de Lima,
Perú*



Capítulo 10

Análisis de los datos cuantitativos



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.

OBJETIVOS DE APRENDIZAJE

Al terminar este capítulo, el alumno será capaz de:

- Revisar el proceso para analizar los datos cuantitativos.
- Reforzar los conocimientos estadísticos fundamentales.
- Comprender las principales pruebas o métodos estadísticos desarrollados, así como sus aplicaciones y la forma de interpretar sus resultados.
- Analizar la interrelación entre distintas pruebas estadísticas.
- 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 son empleados por 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 y no manualmente. Este capítulo se complementa con el capítulo 8, "Análisis estadístico: Pruebas adicionales y análisis multivariado de los datos", del CD anexo.

¿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 nadie lo hace de forma manual, en especial si hay un volumen considerable de datos. Por otra parte, casi en todas 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.

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.

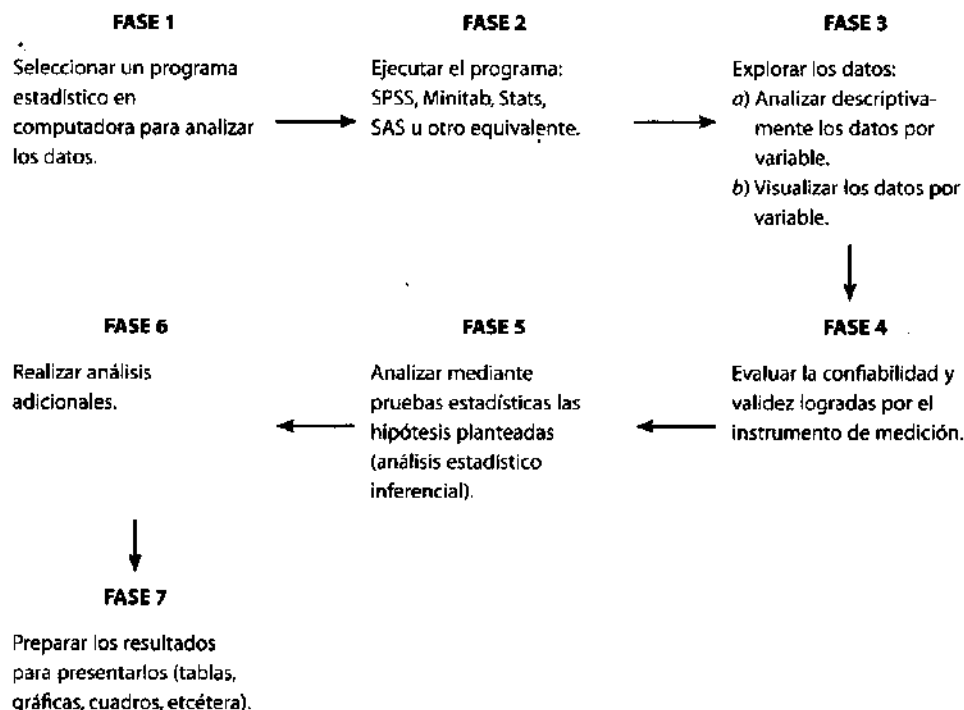


Figura 10.1 Proceso para efectuar análisis estadístico.

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) 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 definir 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—, 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 o Paquete Estadístico para las Ciencias Sociales (SPSS®)

Este programa, desarrollado en la Universidad de Chicago, es uno de los más difundidos. Contiene todos los análisis estadísticos que se describirán en este capítulo y su extensión en el CD anexo.

En América Latina, algunas instituciones educativas tienen versiones antiguas del SPSS; otras, versiones más recientes y en español (en 2005, la última versión era la 14.0). 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 *softwares*, SPSS constantemente se actualiza con versiones nuevas en inglés y español. Asimismo, cada año surgen textos o manuales acordes con las nuevas versiones. Por ello, no vale la pena mencionar alguna referencia, pues será obsoleta cuando el lector revise estas líneas. Lo mejor para mantenerse al día en materia de SPSS es consultar su sitio en internet (<http://www.spss.com/>); o si éste llega a cambiar, con la palabra clave "SPSS" podemos encontrarla en un directorio o mediante un motor de búsqueda como Google, Altavista, o cualquier otro. Para la actualización, las palabras clave serían: "SPSS manuals" (recordemos que para cruzar palabras, éstas tienen que ir entre comillas " ").

Como ya se señaló, SPSS 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). La cual contiene los comandos para operar en la parte superior. En la figura 10.2 se muestra una vista de los datos en SPSS.

En la página de SPSS se puede "bajar" o "descargar" a la computadora una demostración del programa.

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

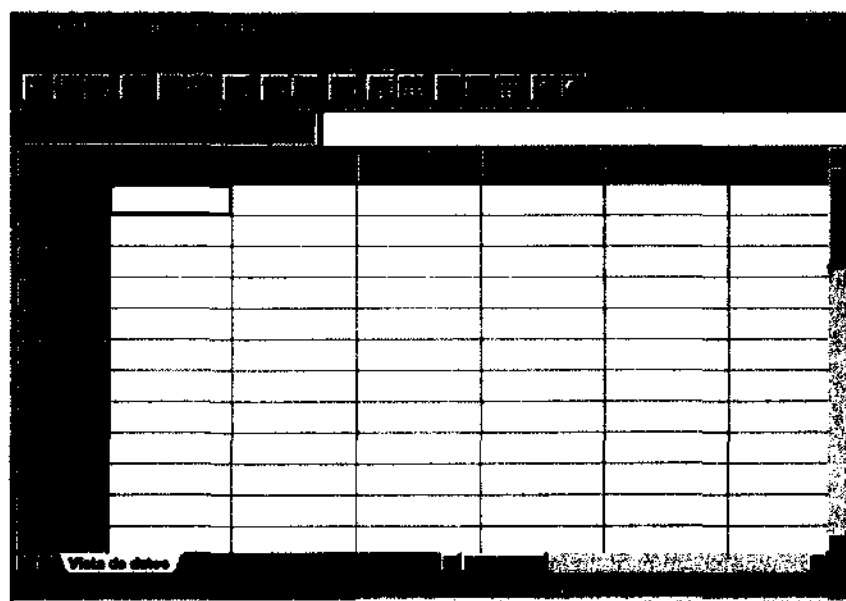


Figura 10.2 Vista de los datos en SPSS.

El paquete SPSS en versión para ambiente Windows 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 que se encuentran en dicho contexto.

File (archivos): 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 regiones).
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 (confiabilidad y escalamiento multidimensional).
11. Pruebas no paramétricas.
12. Respuestas múltiples (escalas).
13. Validación compleja.
14. Series de tiempos.
15. Ecuaciones estructurales y modelamiento matemático.

Graphs (gráficas): 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 (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 y demás elementos de Windows 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 archivo muestra del programa.

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. Para hacer análisis y gráficas, especificar impresora, imprimir, cerrar, entre otras funciones.

Edit (edición): Útil para modificar archivos, buscar datos, copiar, cortar y eliminar celdas, etcétera.

Data (datos): Funciones para dividir la matriz, copiar columnas, eliminar columnas y renglones o filas, establecer rangos, recodificar, cambiar el tipo de datos, desplegar datos, entre otros.

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

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

1. Básicas: Descriptivas, correlación, covarianza, *chi*-cuadrada, prueba *t*...
2. Regresión lineal y múltiple.
3. Análisis de varianza (ANOVA) unidireccional y factorial.
4. DOE (Análisis de factores y de respuestas).
5. Diagramas (de atributos, multivariados, de tiempo) individuales y grupales.
6. Diagramas de dispersión, Pareto, causa-efecto...
7. Confiabilidad.
8. Análisis multivariado: Conglomerados, análisis de factores (validación), análisis discriminante, 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 (Diagramas de caja, fotograma, etcétera).
13. Poder y tamaño de muestra (1-muestra *z*, 1-muestra *t*, 2-muestra *t*, ANOVA y otras).

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

Editr (editar): Mover columnas, redefinir columnas, insertar columnas, buscar, ir a 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, etcétera.

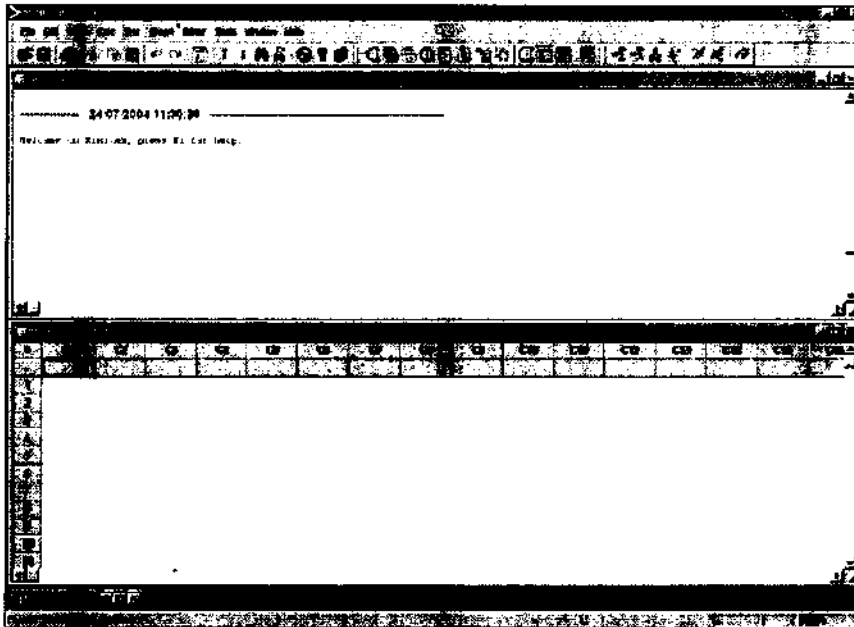


Figura 10.3 Pantalla de Minitab.

Help (ayuda): Cuenta con contenidos de ayuda, cómo utilizar Minitab, comandos, guías y demás elementos de Windows aplicados al paquete. En la figura 10.3 se muestra una vista de la pantalla de Minitab.

Otro programa de análisis sumamente difundido es el SAS (Sistema de Análisis Estadístico). 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 texto 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 se muestra en la figura 10.4 (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 nada para el lector que se inicia en los menesteres de la investigación, pero éstos se irán explicando a lo largo del capítulo.

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 para ser medidas, pero en otras se necesitan varios ítems para tal la finalidad. Cuando sólo se precisa de un ítem, 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.

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 (moral), las preguntas podrían ser las números: 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 a 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

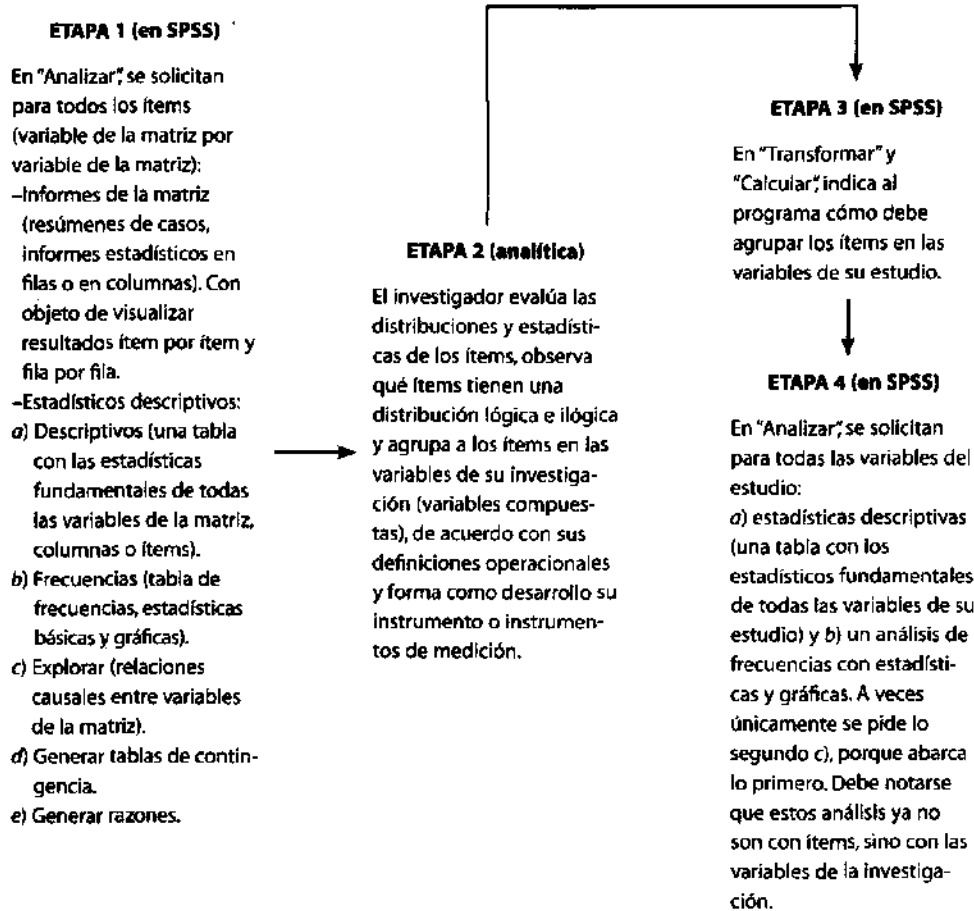


Figura 10.4 Secuencia más común para explorar datos en SPSS.

Tabla 10.2 Ejemplos de variables de investigación y formulación de ítems

Variable con un ítem o variable de la matriz	Variable con tres ítems o variables en la matriz	Variable con cinco ítems o variables en la matriz
¿Asiste a una escuela pública o privada? ☐ Escuela pública. ☐ Escuela privada.	1) ¿En qué medida está usted satisfecho con su superior inmediato? ♣ Sumamente insatisfecho. ♣ Más bien insatisfecho. ♣ Ni insatisfecho, ni satisfecho. ♣ Más bien satisfecho. ♣ Sumamente satisfecho.	1) "En el departamento donde trabajamos mantenemos unidos". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo.

(continúa)

Tabla 10.2 Ejemplos de variables de investigación y formulación de ítems (*continuación*)

Variable con un ítem o variable de la matriz	Variable con tres ítems o variables en la matriz	Variable con cinco ítems o variables en la matriz
Esta variable (tipo de escuela a la que asiste) es medida por una pregunta y ocupa una columna o variable de la matriz.	2) ¿Qué tan satisfecho está usted con el trato que recibe de parte de su superior inmediato? ♣ Sumamente insatisfecho. ♣ Más bien insatisfecho. ♣ Ni insatisfecho, ni satisfecho. ♣ Más bien satisfecho. ♣ Sumamente satisfecho.	2) "La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo.
	3) ¿Qué tan satisfecho está con la orientación que le proporciona su superior inmediato para que usted realice su trabajo? ♣ Sumamente insatisfecho. ♣ Más bien insatisfecho. ♣ Ni insatisfecho, ni satisfecho. ♣ Más bien satisfecho. ♣ Sumamente satisfecho.	3) "En mi departamento nos mantenemos en contacto permanentemente". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo.
	Esta variable (satisfacción con su superior inmediato) es medida por tres preguntas y ocupa tres columnas o variables de la matriz.	4) "En mi departamento nos reunimos con frecuencia para hablar tanto de asuntos de trabajo como de cuestiones personales". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo. 5) "En mi trabajo todos nos llevamos muy bien". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo.
		Esta variable (moral en el grupo de trabajo) es medida por cinco preguntas y ocupa cinco columnas o variables de la matriz.

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*, 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", de este modo, se construye la variable compuesta mediante una expresión numérica. Revisemos un ejemplo:

En el caso de la variable "moral", 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	Variabes de la matriz que corresponden a la variable de la investigación	Ubicación en la matriz
"En el departamento donde trabajo nos mantenemos unidos". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo.	Frase 1 (fr1)	Columna 1
"La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo.	Frase 2 (fr2)	Columna 2
"En mi departamento nos mantenemos en contacto permanentemente". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo.	Frase 3 (fr3)	Columna 3

(continúa)

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

Variable de la investigación: moral	Variables de la matriz que corresponden a la variable de la investigación	Ubicación en la matriz
"En mi departamento nos reunimos con frecuencia para hablar tanto de asuntos de trabajo como de cuestiones personales". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo.	Frase 4 (fr4)	Columna 4
"En mi trabajo todos nos llevamos muy bien". ♣ Totalmente de acuerdo. ♣ De acuerdo. ♣ Ni de acuerdo, ni en desacuerdo. ♣ En desacuerdo. ♣ Totalmente en desacuerdo.	Frase 5 (fr5)	Columna 5

Con la función "Calc", 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:

- El *nivel de medición* de las variables.
- La *manera* como se hayan formulado las *hipótesis*.
- 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 estadística 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 graficaciones 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?

Una **distribución de frecuencias** es un conjunto de puntuaciones ordenadas en sus respectivas categorías. La tabla 10.4 muestra un ejemplo de una distribución de frecuencias.

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

En un estudio entre 200 personas latinas que viven en el estado de California, EE.UU.¹ 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 (en SPSS: <i>prefoe</i>)		
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

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 el tabla 10.6.

¿Qué otros elementos contiene una distribución de frecuencias?

Las distribuciones de frecuencias pueden completarse agregando los porcentajes de casos en cada categoría, los porcentajes válidos (excluyendo los valores perdidos) y los porcentajes acumulados (porcentaje de lo que se va acumulando en cada categoría, desde la más baja hasta la más alta). La tabla 10.7 muestra un ejemplo con los porcentajes en sí, los porcentajes válidos y los acumulados.

El *porcentaje acumulado* constituye lo que aumenta en cada categoría de manera porcentual y progresiva (en orden descendente de aparición de las categorías), tomando en cuenta los *porcentajes válidos*. En la categoría “sí se ha obtenido la cooperación”, se ha acumulado 74.6%. En la categoría “no se ha obtenido la cooperación”, se acumula 78.7% (74.6% de la categoría anterior y 4.1% de la categoría en cuestión). En la última categoría siempre se acumula el total (100%).

Las *columnas porcentaje* y *porcentaje válido* son iguales (mismas cifras o valores) cuando no hay valores perdidos; pero si tenemos valores perdidos, la columna *porcentaje válido* presenta los cálculos sobre el total menos tales valores. En la tabla 10.8 se muestra un ejemplo con valores perdidos en el caso de un estudio exploratorio sobre los motivos de los niños celayenses para elegir su personaje televisivo favorito (García y Hernández Sampieri, 2005).

Al elaborar el reporte de resultados, una distribución se presenta con los elementos más informativos para el lector y la descripción de los resultados o un comentario, tal como se muestra en la tabla 10.9.

¹ Encuesta con 7% de margen de error (University of Southern California y Bendixen & Associates, 2002).

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

En la tabla 10.9 pudieron haberse incluido solamente los porcentajes y eliminarse las frecuencias.

En los comentarios sobre las distribuciones de frecuencias se utilizan frases como “la mitad de los entrevistados prefiere la marca *X*” (con 50%), “poco menos de la mitad” de la población mencionó que votarán por el candidato *X* (por ejemplo, con 48.7%), “casi la tercera parte...” (por ejemplo, con 32.8%), “cuatro de cada diez señoras...” (40%), “solamente uno de cada diez...” (10%), “la mayoría...” (96.7%), etcétera.

Tabla 10.7 Ejemplo de una distribución de frecuencias con todos sus elementos (en SPSS)

Variable: cooperación del personal con el proyecto de calidad de la empresa

Categorías	Códigos	Frecuencias	Porcentaje válido	Porcentaje acumulado
-Sí se ha obtenido la cooperación	1	91	74.6	74.6
-No se ha obtenido la cooperación	2	5	4.1	78.7
-No respondieron	3	26	21.3	100.0
Total		122	100.0	

Tabla 10.8 Ejemplo de tabla con valores perdidos

Motivos de la preferencia de su personaje favorito

		Frecuencias	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Divertidos	142	72.1	73.2	73.2
	Buenos	10	5.1	5.2	78.4
	Tienen poderes	23	11.7	11.9	90.2
	Son fuertes	19	9.6	9.8	100.0
	Total	194	98.5	100.0	
Perdidos	No contestaron	3	1.5		
TOTAL		197	100.0		

Tabla 10.9 Ejemplo de una distribución de frecuencias para presentar a un usuario

¿Se ha obtenido la cooperación del personal para el proyecto de calidad?

Obtención	Núm. de organizaciones	Porcentajes
Sí	91	74.6
No	5	4.1
No respondieron	26	21.3
Total	122	100.0

COMENTARIO. Prácticamente tres cuartas partes de las organizaciones sí han obtenido la cooperación del personal. Llama la atención que poco más de una quinta parte no quiso comprometerse con su respuesta. Las organizaciones que no han logrado la cooperación del personal mencionaron como factores ausentismo, rechazo al cambio y conformismo.

En SPSS se solicitan tales gráficos en: **Anализar** → **Estadísticos descriptivos** → **Frecuencias** → **Gráficos** (ya sea un histograma, de barras, de sectores o "pastel"), además de que se obtienen las distribuciones de frecuencias y estadísticas descriptivas; o bien, directamente en "Gráficos".

¿De qué otra manera pueden presentarse las distribuciones de frecuencias?

Las distribuciones de frecuencias, especialmente cuando utilizamos los porcentajes, pueden presentarse en forma de histogramas o gráficas de otro tipo. Algunos ejemplos se muestran en la figura 10.5.

En la actualidad se dispone de una gran variedad de programas y paquetes computacionales que elaboran cualquier gráfica, a colores, utilizando efectos de movimiento y en tercera dimensión.

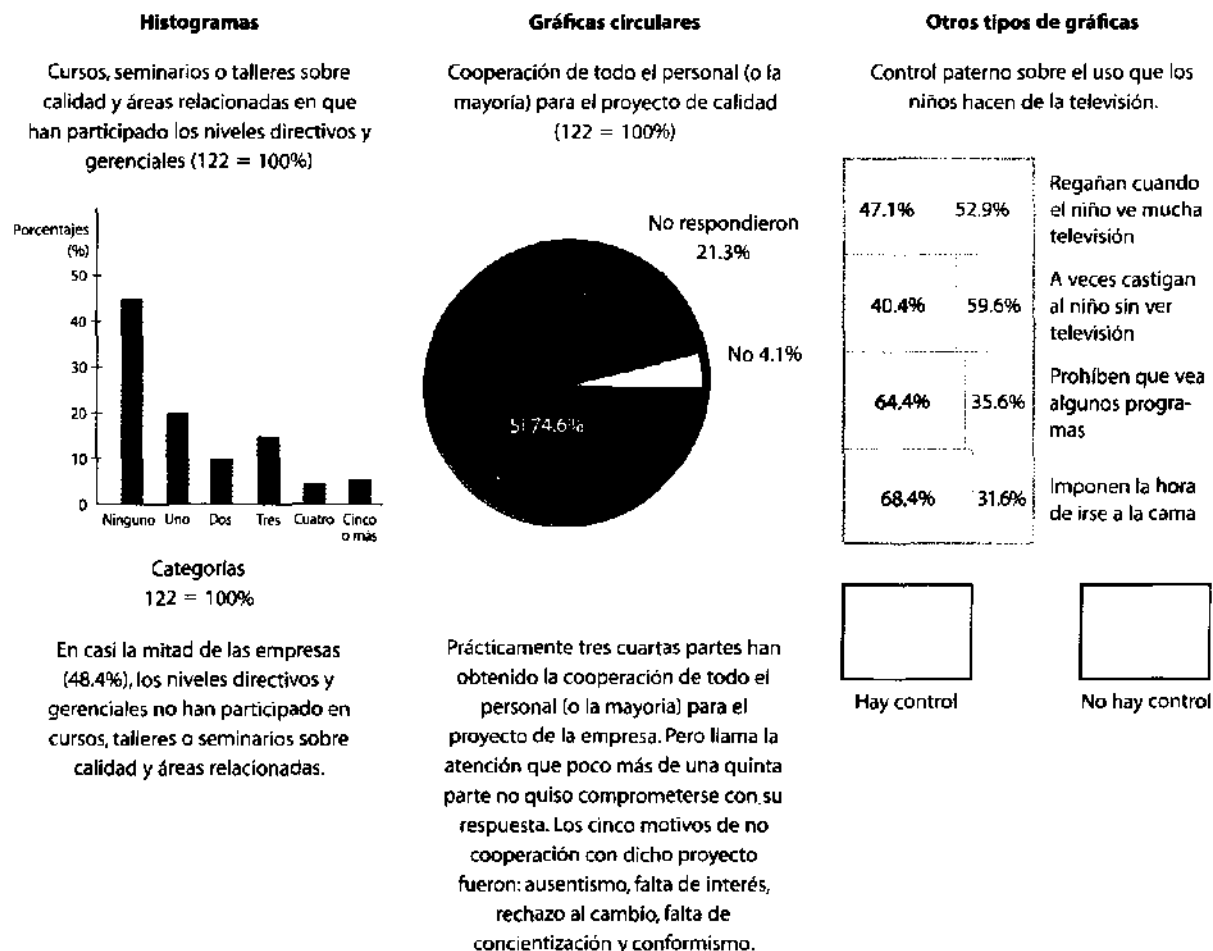


Figura 10.5 Ejemplos de gráficas para presentar distribuciones.

Las distribuciones de frecuencias también se pueden graficar como polígonos de frecuencias

Los **polígonos de frecuencias** relacionan las puntuaciones con sus respectivas frecuencias. Es más bien propio de un nivel de medición por intervalos o razón. Los polígonos se construyen sobre los puntos medios de los intervalos. Por ejemplo, si los intervalos fueran 25-29, 30-34, 35-39, y siguientes; los puntos medios serían 27, 32, 37, etc. SPSS o Minitab realizan esta labor en forma automática.

Polígonos de frecuencias Relacionan las puntuaciones con sus respectivas frecuencias, por medio de gráficas útiles para describir los datos.

Un ejemplo de un polígono de frecuencias se muestra en la figura 10.6.

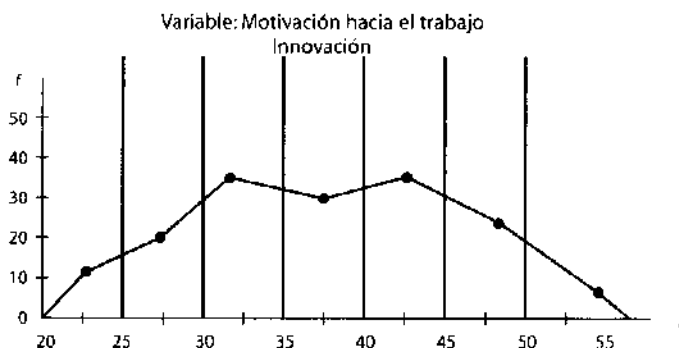


Figura 10.6 Ejemplo de un polígono de frecuencias.

El polígono de frecuencias obedece a la siguiente distribución:

<i>Categorías/intervalos</i>	<i>Frecuencias absolutas</i>
20-24.9	10
25-29.9	20
30-34.9	35
35-39.9	33
40-44.9	36
45-49.9	27
50-54.9	8
TOTAL	169

Los **polígonos de frecuencias** representan curvas útiles para describir los datos. Nos indican hacia dónde se concentran los casos (personas, organizaciones, segmentos de contenido, mediciones de polución, etc.) en la escala de la variable; más adelante se hablará de ello.

En resumen, para cada una de las variables de la investigación se obtiene su distribución de frecuencias y, de ser posible, se grafica y obtiene su polígono de frecuencias correspondiente (en SPSS con los comandos: Gráficos → Líneas → Simple, cuando es un polígono de una variable; cuando en la misma gráfica se desea colocar polígonos con dos o más variables es → Múltiple). En la figura 10.7 se muestra un ejemplo más.

Con respecto a la innovación en la empresa, que es la percepción del apoyo a las iniciativas tendientes a introducir mejoras en la manera como se realiza el trabajo, a nivel organizacional y departamental, la mayoría de los sujetos tienden a estar en altos niveles de la escala.

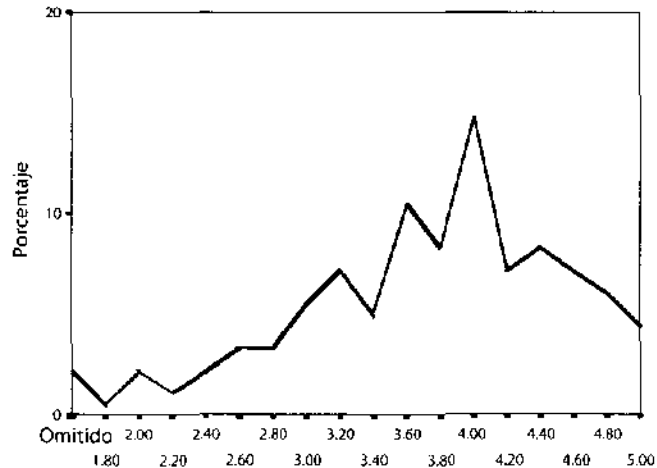


Figura 10.7 Ejemplo de un polígono de frecuencias con la variable innovación.

El polígono puede presentarse con frecuencias como en la figura 10.6 o con porcentajes como con este último ejemplo.

Pero además del polígono de frecuencias, deben calcularse las *medidas de tendencia central* y de *variabilidad o dispersión*.

¿Cuáles son las medidas de tendencia central?

Las **medidas de tendencia central** son puntos en una distribución, los valores medios o centrales de ésta, y nos ayudan a ubicarla dentro de la escala de medición. Las principales medidas de tendencia central son tres: *moda*, *mediana* y *media*. El nivel de medición de la variable determina cuál es la medida de tendencia central apropiada.

La **moda** es la categoría o puntuación que ocurre con mayor frecuencia. En la tabla 10.7, la moda es “1” (sí se ha obtenido la cooperación). Se utiliza con cualquier nivel de medición.

La **mediana** es el valor que divide la distribución por la mitad. Esto es, la mitad de los casos caen por debajo de la mediana y la otra mitad se ubica por encima de la mediana. La mediana refleja la posición intermedia de la distribución. Por ejemplo, si los datos obtenidos fueran:

24 31 35 35 38 43 45 50 57

La mediana es 38, porque deja cuatro casos por encima (43, 45, 50 y 57) y cuatro casos por debajo (35, 35, 31 y 24). Parte a la distribución en dos mitades. En general, para descubrir el caso o la puntuación que constituye la mediana de una distribución, simplemente se aplica la fórmula: $\frac{N + 1}{2}$

Medidas de tendencia central Valores medios o centrales de una distribución que sirven para ubicarla dentro de la escala de medición.

Moda Categoría o puntuación que se presenta con mayor frecuencia.

Si tenemos nueve casos, $\frac{9 + 1}{2}$ entonces buscamos el quinto valor y éste es la mediana. Note que la mediana es el valor observado que se localiza a la mitad de la distribución, no el valor de cinco. La fórmula no nos proporciona directamente el valor de la mediana, sino el número de caso en donde está la mediana.

La **mediana** es una medida de tendencia central propia de los niveles de medición ordinal, por intervalos y de razón. No tiene sentido con variables nominales, porque en este nivel no hay jerarquías ni noción de encima o debajo. Asimismo, la mediana es particularmente útil cuando hay valores extremos en la distribución. No es sensible a éstos. Si tuviéramos los siguientes datos:

24 31 35 35 38 43 45 50 248

la mediana seguiría siendo 38.

Para la interpretación de la media y la mediana, se incluye un artículo al respecto en el siguiente ejemplo.²

EJEMPLO

Interpretación de la mediana

¿Qué edad tiene? Si teme contestar no se preocupe, los perfiles de edad difieren de un país a otro.

Con base en proyecciones sobre la población en 2005 (United Nations Population Fund, 2005), la población mundial para finales de 2007 será de aproximadamente 6 650 millones de habitantes.

El promedio de edad mundial es ligeramente superior a los 26 años (Televisa y Agencia EFE, 2005; Di Santo, 2005). La mediana de edad es similar, más de 26 años (Organización de las Naciones Unidas, 2005), lo que significa que la mitad de los habitantes del globo terrestre sobrepasa esta edad y el otro medio es más joven. Cabe señalar que la mediana varía de un lugar a otro, ya que en los países desarrollados la edad mediana de la población —esto es, la edad que divide a la población en dos partes iguales— ha ido en ascenso constante desde 1950 hasta llegar en el año 2000 a 37.3 años. En los países en desarrollo la edad mediana disminuyó de 21.6 a 19.2 años entre 1950 y 1970; pero a partir de entonces ha registrado un ascenso continuo, hasta alcanzar 25.2 años en 2000. En los países menos desarrollados la edad mediana en 2000 ascendía a 18.5 años. Se estima que a mediados del siglo XXI la edad mediana mundial habrá aumentado en 10 años (a 36 años). Ésta será de 49.9 años en los países desarrollados y de 41.3 años en los países en desarrollo, por lo que la brecha que los separará será sólo de ocho años. En los países menos desarrollados la edad mediana se espera que aumente a 30.1 para el 2050, por lo que todavía será muy inferior a la de los países desarrollados y en desarrollo. Actualmente, el país con la población más joven

² El artículo original es de Leguizamo (1987), pero se actualizó con otras fuentes.

es Yemen, con una edad mediana de 15 años, y el más viejo es Japón, con una edad mediana de 41 años.

Buena noticia para el actual ciudadano global medio, porque parece ser que se encuentra en la situación de envejecer más lentamente.

$$\bar{X}$$

La **media** es la medida de tendencia central más utilizada y puede definirse como el promedio aritmético de una distribución. Se simboliza como $\bar{X}$, y es la suma de todos los valores dividida entre el número de casos. Es una medida solamente aplicable a mediciones por intervalos o de razón. Carece de sentido para variables medidas en un nivel nominal u ordinal. Su fórmula es:

Media Es el promedio aritmético de una distribución y es la medida de tendencia central más utilizada.

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_k}{N}$$

Por ejemplo, si tuviéramos las siguientes puntuaciones:

8 7 6 4 3 2 6 9 8

La media sería igual a:

$$\bar{X} = \frac{8 + 7 + 6 + 4 + 3 + 2 + 6 + 9 + 8}{9} = 5.888$$

La fórmula simplificada de la media es:

$$\bar{X} = \frac{\Sigma X}{N}$$

El símbolo " Σ " indica que debe efectuarse una sumatoria, X es el símbolo de una puntuación y N es el número total de casos o puntuaciones. En nuestro ejemplo:

$$\bar{X} = \frac{53}{9} = 5.888$$

La media sí es sensible a valores extremos. Si tuviéramos las siguientes puntuaciones:

8 7 6 4 3 2 6 9 20

La media sería:

$$\bar{X} = \frac{65}{9} = 7.22$$

¿Cuáles son las medidas de la variabilidad?

Las **medidas de la variabilidad** indican la dispersión de los datos en la escala de medición y responden a la pregunta: ¿dónde están diseminadas las puntuaciones o los valores obtenidos? Las medidas de tendencia central son valores en una distribución y las medidas de la variabilidad son intervalos que designan distancias o un número de unidades en la escala de medición. Las medidas de la variabilidad más utilizadas son *rango*, *desviación estándar* y *varianza*.

El **rango**, también llamado *recorrido*, es la diferencia entre la puntuación mayor y la puntuación menor, e indica el número de unidades en la escala de medición que se necesitan para incluir los valores máximo y mínimo. Se calcula así: $X_M - X_m$ (puntuación mayor, menos puntuación menor). Si tenemos los siguientes valores:

17 18 20 20 24 28 28 30 33

El rango será: $33 - 17 = 16$.

Cuanto *más grande sea el rango, mayor será la dispersión de los datos* de una distribución.

La **desviación estándar** o típica es el promedio de desviación de las puntuaciones con respecto a la media. Esta medida se expresa en las unidades originales de medición de la distribución. Se interpreta en relación con la media. Cuanto mayor sea la dispersión de los datos alrededor de la media, mayor será la desviación estándar. Se simboliza con: s o la sigma minúscula σ y su fórmula esencial es:

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{N}}$$

Esto es, la desviación de cada puntuación respecto a la media se eleva al cuadrado, se suman todas las desviaciones cuadradas, se divide entre el número total de puntuaciones, y a esta división se le saca raíz cuadrada.

La desviación estándar se interpreta como *cuánto se desvía, en promedio, de la media un conjunto de puntuaciones*.

Supongamos que un investigador obtuvo para su muestra una media (promedio) de ingreso familiar anual de \$6 000 y una desviación estándar de \$1 000. La interpretación es que los ingresos familiares de la muestra se desvían, en promedio, mil unidades monetarias respecto a la media.

La desviación estándar sólo se utiliza en variables medidas por intervalos o de razón.

Medidas de la variabilidad Son intervalos que indican la dispersión de los datos en la escala de medición.

Rango Indica la extensión total de los datos en la escala.

Desviación estándar Promedio de desviación de las puntuaciones con respecto a la media que se expresa en las unidades originales de medición de la distribución.

La varianza

La **varianza** es la desviación estándar elevada al cuadrado y se simboliza s^2 . Es un concepto estadístico muy importante, ya que muchas de las pruebas cuantitativas se fundamentan en él. Diversos métodos estadísticos parten de la descomposición de la varianza. Sin embargo, con fines descriptivos se utiliza preferentemente la desviación estándar.

Varianza Se utiliza en análisis inferenciales.

¿Cómo se interpretan las medidas de tendencia central y de la variabilidad?

Cabe destacar que al describir nuestros datos, respecto a cada *variable del estudio*, interpretamos las medidas de tendencia central y de la variabilidad en conjunto, no aisladamente. Consideramos todos los valores. Para interpretarlos, lo primero que hacemos es tomar en cuenta el rango potencial de la escala. Supongamos que aplicamos una escala de actitudes del tipo Likert para medir la “actitud hacia el presidente” de una nación (digamos que la escala tuviera 18 ítems y se promediaran sus valores). El rango potencial es de uno a cinco (ver figura 10.8).

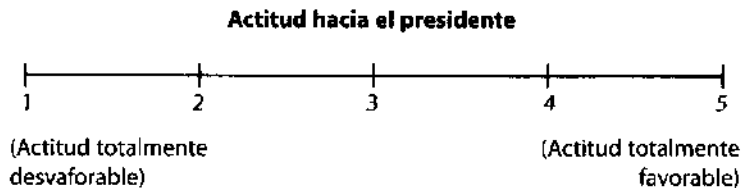


Figura 10.8 Ejemplo de escala con rango potencial.

Si obtuviéramos los siguientes resultados:

Variable: actitud hacia el presidente

Moda: 4.0

Mediana: 3.9

Media ($\bar{X}$): 4.2

Desviación estándar: 0.7

Puntuación más alta observada (máximo): 5.0

Puntuación más baja observada (mínimo): 2.0

Rango: 3

Podríamos hacer la siguiente interpretación descriptiva: la actitud hacia el presidente es favorable. La categoría que más se repitió fue 4 (favorable). Cincuenta por ciento de los sujetos está por encima del valor 3.9 y el restante 50% se sitúa por debajo de este valor (mediana). En promedio, los sujetos se ubican en 4.2 (favorable). Asimismo, se desvían de 4.2, en promedio, 0.7 unidades de la escala. Ninguna persona calificó al presidente de manera muy desfavorable (no hay “1”). Las puntuaciones tienden a ubicarse en valores medios o elevados.

En cambio, si los resultados fueran:

Variable: actitud hacia el presidente

Moda: 1

Mediana: 1.5

Media ($\bar{X}$): 1.3

Desviación estándar: 0.4

Varianza: 0.16

Máximo: 3.0

Mínimo: 1.0

Rango: 2.0

La interpretación es que la actitud hacia el presidente es muy desfavorable. En la figura 10.9 vemos gráficamente la comparación de resultados.

La variabilidad también es menor en el caso de la actitud muy desfavorable (los datos se encuentran menos dispersos).

En la tabla 10.10 se presenta otro ejemplo de interpretación con una prueba de motivación intrínseca aplicada a 60 participantes de un experimento (Hernández Sampieri y Cortés, 1982). La escala tiene 17 ítems (con cinco opciones cada uno, uno a cinco) y mide la motivación intrínseca al ejecutar una tarea.

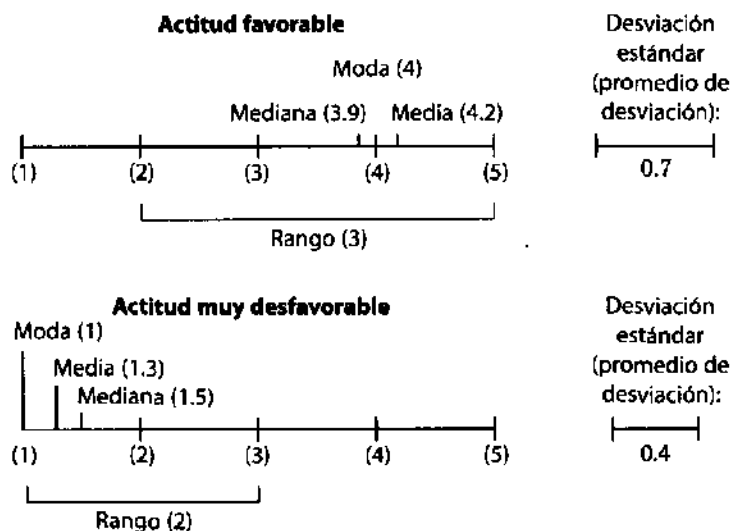


Figura 10.9 Ejemplo de interpretación gráfica de las estadísticas descriptivas.

El nivel de motivación intrínseca exhibido por los sujetos tiende a ser elevado, como lo indican los resultados de la escala. El rango real de la escala iba de 17 a 85. El rango resultante para esta investigación varió de 40 a 81. Por lo tanto, es evidente que los sujetos se inclinaron hacia valores elevados en la medida de motivación intrínseca. Además, la media de los participantes es

Tabla 10.10 Ejemplo de interpretación de una distribución de frecuencias

Valores registrados en la escala de motivación intrínseca	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
40	1	1.7	1.7	1.7
44	1	1.7	1.7	3.3
48	1	1.7	1.7	5.0
51	1	1.7	1.7	6.7
52	2	3.3	3.3	10.0
56	2	3.3	3.3	13.3
58	1	1.7	1.7	15.0
59	1	1.7	1.7	16.7
60	2	3.3	3.3	20.0
61	4	6.7	6.7	26.7
63	2	3.3	3.3	30.0
64	2	3.3	3.3	33.3
65	3	5.0	5.0	38.3
66	2	3.3	3.3	41.7
67	4	6.7	6.7	48.3
68	3	5.0	5.0	53.3
69	1	1.7	1.7	55.0
70	4	6.7	6.7	61.7
71	3	5.0	5.0	66.7
72	4	6.7	6.7	73.3
73	3	5.0	5.0	78.3
74	2	3.3	3.3	81.7
75	1	1.7	1.7	83.3
76	1	1.7	1.7	85.0
77	2	3.3	3.3	88.3
78	1	1.7	1.7	90.0
79	2	3.3	3.3	93.3
80	2	3.3	3.3	96.7
81	2	3.3	3.3	100.0
Total	60	100.0	100.0	

Media = 66.883

Moda = 61.000

Curtosis = .587

Mínimo = 40.000

E.E. = 1.176

s = 9.112

Asimetría = -.775

Máximo = 81.000

Mediana = 67.833

Varianza = 83.020

Rango = 41.000

Sumatoria = 4 013.000

de 66.9 y la mediana de 67.8, lo cual confirma la tendencia de la muestra hacia valores altos de la escala. A pesar de que la dispersión de las puntuaciones de los sujetos es considerable (la desviación estándar es igual a 9.1 y el rango es de 41), esta dispersión se manifiesta en el área más elevada de la escala. Veámoslo gráficamente en la figura 10.10:

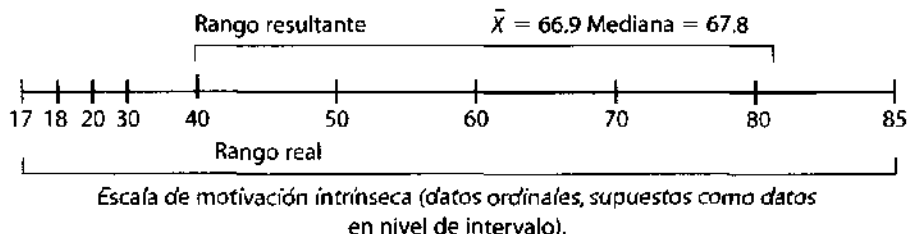


Figura 10.10 Gráfica del rango de la tabla 10.10.

En resumen, la tarea resultó intrínsecamente motivante para la mayoría de los sujetos; sólo que para algunos resultó muy motivante; para otros, relativamente motivante; y para los demás, medianamente motivante. Esto es, que la tendencia general es hacia valores superiores (observamos la columna de frecuencias acumuladas y notamos que 80% obtuvo puntuaciones mayores a 60, mientras que 20% se ubica por debajo).

Ahora bien, ¿qué significa un alto nivel de motivación intrínseca exhibido con respecto a una tarea? Significa que la tarea fue percibida como atractiva, interesante, divertida y categorizada como una experiencia agradable. Asimismo, implica que los sujetos, al ejecutarla, derivaron de ella sentimientos de satisfacción, goce y realización personal. Por lo general, quien se encuentra intrínsecamente motivado hacia una labor, disfrutará la ejecución de ésta, ya que obtendrá de la labor *per se* recompensas internas, como sentimientos de logro y autorrealización. Además de ser absorbido por el desarrollo de la tarea y, al tener un buen desempeño, la opinión de sí mismo mejorará o se verá reforzada.

¿Hay alguna otra estadística descriptiva?

Sí, la *asimetría* y la *curtosis*. Los *polígonos de frecuencia* suelen representarse como *curvas* (figura 10.11) para que puedan analizarse en términos de probabilidad y visualizar su grado de dispersión. De hecho, en realidad son curvas. Los dos elementos mencionados son esenciales para estas curvas o polígonos de frecuencias.

La **asimetría** es una estadística necesaria para conocer cuánto se parece nuestra distribución a una distribución teórica llamada *curva normal* (la cual se representa en la figura 10.11) y constituye un indicador del lado de la curva donde se agrupan las frecuencias. Si es cero (asimetría = 0), la curva o distribución es simétrica. Cuando es positiva, quiere decir que hay más valores agrupados hacia la izquierda de la curva (por debajo de la media). Cuando es negativa, significa que los valores tienden a agruparse hacia la derecha de la curva (por encima de la media).

Asimetría Estadística que se usa para conocer cuánto se parece una distribución a la distribución teórica llamada *curva normal* y que constituye un indicador del lado de la curva en el que las frecuencias se agrupan más.

La **curtosis** es un indicador de lo plana o “picuda” que es una curva. Cuando es cero ($\text{curtosis} = 0$), significa que puede tratarse de una *curva normal*. Si es positiva, quiere decir que la curva, la distribución o el polígono es más “picuda(o)” o elevada(o). Si la curtosis es negativa, indica que es más plana la curva.

Curtosis Indicador de lo plana o *picuda* que es una curva.

La asimetría y la curtosis requieren mínimo de un nivel de medición por intervalos. En la figura 10.11 se muestran ejemplos de curvas con su interpretación.



Distribución simétrica ($\text{asimetría} = 0$), con curtosis positiva, y una desviación estándar y varianza medias.



Distribución con asimetría negativa, curtosis positiva, y desviación estándar y varianza mayores.



Distribución con asimetría positiva, curtosis negativa, y desviación estándar y varianza considerables.



Distribución con asimetría negativa, curtosis positiva, y desviación estándar y varianza menores.



Distribución simétrica, curtosis positiva, y una desviación estándar y varianza bajas.



Curva normal, $\text{curtosis} = 0$, $\text{asimetría} = 0$, y desviación estándar y varianza promedios.

Figura 10.11 Ejemplos de curvas o distribuciones y su interpretación.

¿Cómo se traducen las estadísticas descriptivas al inglés?

Algunos programas y paquetes estadísticos computacionales pueden realizar el cálculo de las estadísticas descriptivas, cuyos resultados aparecen junto al nombre respectivo de éstas, muchas veces en inglés.

A continuación se indican las diferentes estadísticas y su equivalente en inglés.

<i>Estadística</i>	<i>Equivalente en inglés</i>
— Moda	— <i>Mode</i>
— Mediana	— <i>Median</i>
— Media	— <i>Mean</i>
— Desviación estándar	— <i>Standard deviation</i>
— Varianza	— <i>Variance</i>
— Máximo	— <i>Maximum</i>
— Mínimo	— <i>Minimum</i>
— Rango	— <i>Range</i>
— Asimetría	— <i>Skewness</i>
— Curtosis	— <i>Kurtosis</i>

Nota final

Debe recordarse que en una investigación se obtiene una distribución de frecuencias y se calculan las estadísticas descriptivas para cada *variable*, las que se necesitan de acuerdo con los propósitos de la investigación y los niveles de medición.

EJEMPLO

Hernández Sampieri (2005), en su investigación sobre el clima organizacional, obtuvo las siguientes estadísticas fundamentales de sus variables en una de las muestras:

	N	Mínimo	Máximo	Media	Desviación estándar
Moral	390	1.00	5.00	3.3818	.91905
Dirección	393	1.00	5.00	2.7904	1.08775
Innovación	396	1.00	5.00	3.4621	.91185
Identificación	383	1.00	5.00	3.6584	.91283
Comunicación	397	1.00	5.00	3.2519	.87446
Desempeño	403	1.00	5.00	3.6402	.86793
Motivación intrínseca	401	2.00	5.00	3.9111	.73900
Autonomía	395	1.00	5.00	3.2025	.85466
Satisfacción	399	1.00	5.00	3.7249	.90591
Liderazgo	392	1.00	5.00	3.4532	1.10019
Visión	391	1.00	5.00	3.7341	.89206
Recompensas	381	1.00	5.00	2.4528	1.14364

Notas: Todas las variables son compuestas (integradas de varios ítems). La columna "N" representa el número de casos válidos para cada variable. El N total de la muestra es de 412, pero como podemos ver en la tabla, el número de casos es distinto en las diferentes variables, porque SPSS elimina de toda la variable a los casos que no hayan respondido a un ítem o más ítems.

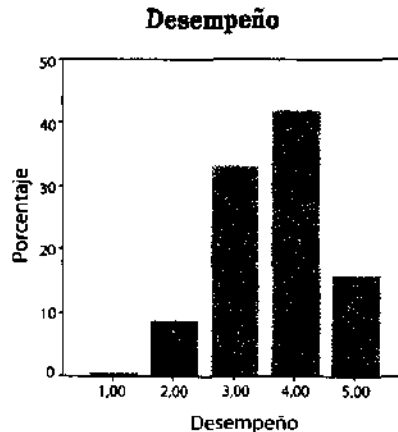
Posteriormente, obtuvo las tablas y distribuciones de frecuencias de todas sus 12 variables. De las cuales solamente incluimos la variable "desempeño" por cuestiones de espacio.

Desempeño

	Valores	Frecuencia	Porcentaje válido	Porcentaje acumulado
	1	2	0.5	0.5
	2	35	8.7	9.2
	3	133	33.0	42.2
	4	169	41.9	84.1
	5	64	15.9	100.0
Total		403	100.0	

N = 420

Perdidos = 17



Puntuaciones z

Las **puntuaciones z** son transformaciones que se pueden hacer a los valores o las puntuaciones obtenidas, con el propósito de analizar su distancia respecto a la media, en unidades de desviación estándar. Una puntuación z nos indica la dirección y el grado en que un valor individual obtenido se aleja de la media, en una escala de unidades de desviación estándar. Como mencionan Nie *et al.* (1975), las puntuaciones z son el método más comúnmente utilizado para estandarizar la escala de una variable medida en un nivel por intervalos.

Su fórmula es:

$$z = \frac{X - \bar{X}}{s}$$

donde *X* es la puntuación o el valor a transformar, $\bar{X}$ es la media de la distribución y *s* la desviación estándar de ésta. El resultado *z* es la puntuación transformada en unidades de desviación estándar.

Puntuación z Medida que indica la dirección y el grado en que un valor individual se aleja de la media, en una escala de unidades de desviación estándar.

Supongamos que en una distribución de frecuencias obtuvimos una media de 60 y una desviación estándar de 10, y deseamos comparar a una puntuación de "50" con el resto de la distribución. Entonces, transformamos esta puntuación o tal valor en una puntuación z . Tenemos que:

$$X = 50$$

$$\bar{X} = 60$$

$$s = 10$$

La puntuación z correspondiente a un valor de "50" es:

$$z = \frac{50 - 60}{10} = -1.00$$

Podemos decir que el valor "50" se localiza a una desviación estándar por debajo de la media de la distribución (el valor "30" está a tres desviaciones estándar por debajo de la media).

Estandarizar los valores permite comparar puntuaciones de dos distribuciones diferentes (la forma de medición es la misma, aunque se trata de distribuciones distintas). Por ejemplo, podemos comparar una distribución obtenida en una preprueba con otra obtenida en una posprueba (en un contexto experimental). Supongamos que se trata de un estímulo que incrementa la productividad. Un trabajador obtuvo en la preprueba una productividad de 130 (la media del grupo fue de 122.5 y la desviación estándar de 10). Y en la posprueba obtuvo 135 (la media del grupo fue de 140 y la desviación estándar de 9.8). ¿Mejoró la productividad del trabajador? En apariencia la mejoría no es considerable. Sin transformar las dos calificaciones en puntuaciones z , no es posible asegurarlo porque los valores no pertenecen a la misma distribución. Entonces transformamos ambos valores a puntuaciones z , los pasamos a una escala común donde la comparación es válida. El valor de 130 en productividad en términos de unidades de desviación estándar es igual a:

$$z = \frac{130 - 122.5}{10.0} = 0.75$$

Y el valor de 135 corresponde a una puntuación z de:

$$z = \frac{135 - 140}{9.8} = -0.51$$

Como observamos, en términos absolutos 135 es una mejor puntuación que 130, pero no en términos relativos (en relación con sus respectivas distribuciones).

La distribución de puntuaciones z no cambia la forma de la distribución original, pero sí modifica las unidades originales a "unidades de desviación estándar" (Wright, 1979). La distribución de puntuaciones z tiene una media de 0 (cero) y una desviación estándar de 1 (uno). La figura 10.12 muestra la distribución de puntuaciones z .

Las puntuaciones z también sirven para comparar mediciones de distintas pruebas o escalas aplicadas a los mismos participantes (los valores obtenidos en cada escala se transforman a puntuaciones z y se comparan). No debe olvidarse que los elementos de la fórmula específicamente son la media y la desviación estándar que corresponden al valor a transformar (de su misma distribución). También, las puntuaciones z sirven para analizar distancias, entre puntuaciones de una misma distribución y áreas de la curva que abarcan tales distancias, o para sopesar el desempeño de un grupo de sujetos en varias pruebas. Las puntuaciones z son un elemento descriptivo adicional que se agrega para analizar nuestros datos.

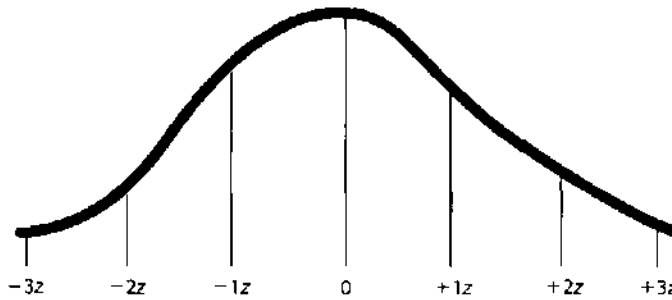


Figura 10.12 Distribución de puntuaciones z .

Razones y tasas

Una **razón** es la relación entre dos categorías. Por ejemplo:

Categorías	Frecuencia
Masculino	60
Femenino	30

La razón de hombres a mujeres es de $\frac{60}{30} = 2$. Es decir, por cada dos hombres hay una mujer.

Una **tasa** es la relación entre el número de casos, frecuencias o eventos de una categoría y el número total de observaciones, multiplicada por un múltiplo de 10, generalmente 100 o 1 000. La fórmula es:

Tasa Es la relación entre el número de casos de una categoría y el número total de observaciones.

$$\text{Tasa} = \frac{\text{Número de eventos durante un periodo}}{\text{Número total de eventos posibles}} \times 100 \text{ o } 1\,000$$

$$\text{Ejemplo: } \frac{\text{Número de nacidos vivos en la ciudad}}{\text{Número de habitantes en la ciudad}} \times 1\,000$$

$$\text{Tasa de nacidos vivos en Santa Lucía: } \frac{10\,000}{300\,000} \times 1\,000 = 33.33$$

Es decir, hay 33.33 nacidos vivos por cada 1 000 habitantes en Santa Lucía.

Ahora bien, hemos analizado descriptivamente los datos por *variable del estudio* y los visualizamos gráficamente. En caso de que alguna distribución resulte ilógica, debemos cuestionarnos si la variable debe ser excluida, sea por errores del instrumento de medición o, en la recolección de los datos, ya que la codificación puede ser verificada. Por ejemplo, supongamos que nos encontramos un porcentaje alto de valores perdidos (de 20%),³ debemos preguntarnos: ¿por qué

³ Un porcentaje de valores perdidos (*missing data*) no debe ser mayor de 15%, no es razonable (Creswell, 2005). Cuando tenemos valores perdidos, podemos ignorarlos o sustituirlos por el valor promedio obtenido del total de puntuaciones válidas, esto lo hacen muchos programas de análisis si así lo deseamos.

tantos participantes no respondieron o contestaron erróneamente? O, al medir la satisfacción laboral, resulta que 90% se encuentra “sumamente satisfecho” (¿es lógico?); u otro caso sería que, en ingresos anuales el promedio fuera de 15 000 dólares por familia (¿resulta creíble en tal municipio?). La tarea es revisar la información descriptiva de todas las variables.

Ahora, debemos demostrar la confiabilidad y validez de nuestro instrumento, sobre la base de los datos recolectados.

... Paso 4: evaluar la confiabilidad y validez lograda por el instrumento de medición

Determinamos las puntuaciones o valores obtenidos por los casos (participantes, objetos, etc.) en cada ítem, tomado individualmente. Algunos ítems constituyen variables (o dicho al revés, hay variables con un solo ítem). Otros ítems, agrupados, miden una misma variable y deben constituir una escala para poder juntarse o sumarse (tener las mismas categorías o alternativas de respuesta), ahora debemos demostrar que tales escalas fueron confiables y válidas en la investigación.

La confiabilidad

En el capítulo anterior se mencionó que la confiabilidad (fiabilidad o congruencia interna) de las escalas se calcula mediante diversos métodos:

1. Medida de estabilidad (confiabilidad por *test-retest*). Que se calcula aplicando a los participantes la misma prueba dos veces y luego obteniendo un coeficiente de correlación entre las puntuaciones de ambas aplicaciones (el tema de relaciones estadísticas se revisará más adelante, por ahora veremos que su interpretación es similar a los demás métodos). Esto se representaría en la figura 10.13:

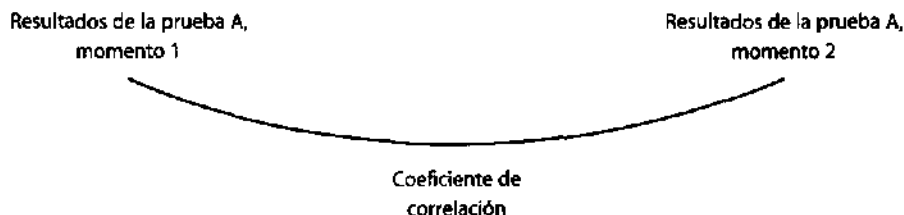


Figura 10.13 Medida de estabilidad.

2. Método de formas alternativas o paralelas. Que se calcula a través de un coeficiente de correlación entre los resultados de dos pruebas supuestamente equivalentes. Lo cual se puede visualizar en la figura (10.14):



Figura 10.14 Método de formas alternativas o paralelas.

3. Método de mitades partidas, que se calcula por medio de un coeficiente de correlación entre las puntuaciones de las mitades del instrumento (se correlacionan los resultados de una mitad del instrumento con los resultados de la otra mitad, aparentemente equivalente). Algo así, como lo esquematizado en la figura 10.15:



Figura 10.15 Método de mitades partidas.

4. Medidas de coherencia interna. Coeficientes de confiabilidad *alfa* de Cronbach (α) y los coeficientes KR-20 y KR-21.

Todos estos coeficientes oscilan entre 0 y 1, donde un coeficiente de 0 significa nula confiabilidad y 1 representa un máximo de confiabilidad (confiabilidad total). Incluso en el capítulo 9 se visualizó este continuo. No hay una regla que indique: a partir de este valor no hay fiabilidad del instrumento. Más bien, el investigador calcula su valor, lo reporta y lo somete a escrutinio de los usuarios del estudio u otros investigadores. Pero podemos decir —de manera más o menos general— que si obtengo 0.25 en la correlación o coeficiente, esto indica baja confiabilidad; si el resultado es 0.50, la fiabilidad es media o regular. En cambio, si supera el 0.75 es aceptable, y si es mayor a 0.90 es elevada, para tomar muy en cuenta.

Con respecto a los métodos basados en coeficientes de correlación, rogamos al lector se forme una idea más clara después de revisar el apartado de correlación que se presenta más adelante en este capítulo. Pero sí hay una consideración importante que hacer ahora. El coeficiente que elijamos para determinar la confiabilidad debe ser apropiado al nivel de medición de la escala de nuestra variable (por ejemplo, si la escala de mi variable es por intervalos, puedo utilizar el coeficiente de correlación de Pearson; pero si es ordinal podré utilizar el coeficiente de Spearman o de Kendall; y si es nominal, otros coeficientes). *Alfa* trabaja con variables de intervalos o de razón y KR-20 y KR-21 con ítems dicotómicos. El cálculo del coeficiente *alfa* se incluye en el capítulo 8 del CD anexo.

En la tabla 10.11 se presentan ejemplos de estudios con su respectiva confiabilidad.

Tabla 10.11 Ejemplos de confiabilidad

Investigación	Instrumento	Métodos de cálculo y resultados	Comentario
Estudio para medir el auto-conocimiento acerca del abuso sexual infantil (Susslawky y Wurtele, 1986).	Los autores desarrollaron el Personal Safety Questionnaire (PSQ), que consta de 13 reactivos. Mide actitudes y conocimiento acerca del abuso en niños y niñas de edad preescolar y de los primeros grados de primaria.	1) <i>Test-retest</i> (estabilidad) con una semana de diferencia entre las aplicaciones y una correlación de 0.64 y 2) una coherencia interna de $r = 0.78$ usando la fórmula Kuder-Richardson (KR- 20).	La confiabilidad es aceptable (0.78) y la correlación entre ambas administraciones media alta.
Evaluación de los conocimientos, opiniones, experiencias y acciones en torno al abuso sexual infantil (Kolko, Moser, Litz y Hughes, 1987).	Escala cognitiva de nueve ítems para infantes en edades preescolares y primeros grados básicos.	1) Coherencia interna α de 0.34.	Confiabilidad baja que demuestra incongruencia, atribuida por los autores a lo corto de la escala (pocos ítems).
Investigación para validar una herramienta que mide el conocimiento acerca del abuso sexual infantil (Hazzard, Webb, Kleemeier, Anger y y Pohl (1991).	Instrumento denominado: What I Know About Touching Scale, usada con infantes de 3° y 4° de primaria.	1) Coherencia interna α Cronbach de 0.75 y 2) confiabilidad por <i>test-retest</i> de 0.77 (con dos semanas de diferencia entre las administraciones).	Fiabilidad aceptable.
Estudio sobre la repercusión que tiene la ansiedad generada por las actividades académicas en el desempeño escolar (Suárez Gallardo, 2004).	Dos escalas de 25 ítems. Una tipo Likert para medir la ansiedad sobre actividades académicas y la otra (también Likert) para el desempeño escolar.	1) El valor de la confiabilidad para la escala de ansiedad, al aplicar una prueba α de Cronbach, fue de 0.916; en tanto que el valor de confiabilidad para la escala de desempeño escolar, al emplear dicha prueba, fue de 0.93.	Las dos mediciones (de la ansiedad generada por las actividades académicas y la del desempeño escolar), indican una estabilidad muy alta.
Desarrollo y validación de una escala autoaplicable para medir la satisfacción sexual en varones y mujeres de México (Álvarez Gayou, Millán y Espinosa, 2005).	Un inventario para medir la satisfacción sexual que está integrado por 29 reactivos y fue administrado a una muestra de 760 personas, de ambos géneros, cuyas edades fluctuaron entre los 16 y 65 años.	La confiabilidad del inventario establecida al aplicar una prueba α Cronbach fue de 0.92.	El valor α indica una estabilidad muy elevada.

(continúa)

Tabla 10.11 Ejemplos de confiabilidad (*continuación*)

Investigación	Instrumento	Métodos de cálculo y resultados	Comentario
Validación de un instrumento para medir el sentido de vida de acuerdo con el pensamiento de V. Frankl (Núñez, 2001).	Una escala tipo Likert con 99 ítems (tercera versión), que mide las siguientes dimensiones: 1) autorrealización 2) experiencia 3) propio bienestar 4) intereses y pasatiempos 5) interés social 6) necesidad vital 7) ocupación principal 8) familia 9) ideales Se administró a 80 profesores.	Se obtuvo un coeficiente <i>alfa</i> igual a 0.96.	Fiabilidad sumamente alta.

Como podemos observar en la tabla 10.11, entre más información se proporcione sobre la confiabilidad, el lector se forma una idea más clara sobre su cálculo y las condiciones en que se demostró. Es indispensable incluir las dimensiones de la variable medida, el tamaño de muestra y el método utilizado. Una cuestión importante es que regularmente los coeficientes son sensibles al número de ítems o reactivos, entre más agreguemos, el valor del coeficiente será más elevado. Insistimos en que el coeficiente *alfa* es para intervalos y los coeficientes Kuder Richardson para ítems dicotómicos (por ejemplo: sí-no). Estos últimos se usan en el método de “mitades partidas”, aunque —como señala Creswell (2005), se confía en la mitad de la información del instrumento, por lo que conviene agregar el cálculo de “profecía” Spearman-Brown.

Además de calcular un coeficiente de correlación y/o un coeficiente de coherencia entre los ítems del instrumento, es conveniente calcular la correlación ítem-escala completa. Ésta representa la vinculación de cada reactivo con toda la escala. Habrá tantas correlaciones como ítems contenga el instrumento. Corbetta (2003, p. 237) lo ejemplifica adecuadamente de la siguiente manera: Si estamos midiendo el autoritarismo, es lógico pensar que, quien alcanza altas puntuaciones en esta variable en toda la escala (es muy autoritaria), habrá de tener puntuaciones elevadas en todos los ítems que la conforman. Pero si uno de los reactivos sistemáticamente (en un número considerable de sujetos) presenta valores contradictorios con respecto a la escala total, podemos concluir que ese ítem no funciona adecuadamente (contradice a los demás reactivos). Los ítems que alcancen coeficientes de correlación bajos con la escala, tal vez deban analizarse y eventualmente, eliminarse.

Asimismo, cada uno de los reactivos puede ser evaluado en su capacidad de discriminación mediante la prueba *t* de Student (paramétrica). Se consideran dos grupos, el primero integrado por 25% de los casos con los puntajes más altos obtenidos en el ítem y el otro grupo compuesto por 25% de los casos con los puntajes más bajos. Los ítems cuya prueba no resulte significativa serán reconsiderados.

Los conceptos estadísticos aquí vertidos tendrán mayor sentido, una vez que se revisen (lo cual se hará más adelante en este capítulo).

La validez

Ya se comentó en el capítulo anterior que la evidencia sobre la validez del contenido se obtiene mediante las opiniones de expertos y al asegurarse que las dimensiones medidas por el instrumento sean representativas del universo o dominio de dimensiones de la(s) variable(s) de interés (a veces mediante un muestreo aleatorio simple). La evidencia de la validez de criterio se produce al correlacionar las puntuaciones de los participantes, obtenidas por medio del instrumento, con sus valores logrados en el criterio. Recordemos que una correlación implica asociar puntuaciones obtenidas por la muestra en dos o más variables.

Por ejemplo, Núñez (2001), además de aplicar su instrumento sobre el sentido de vida, administró otras dos pruebas que supuestamente miden variables similares: el PIL (propósito de vida) y el Logo-test de Elizabeth Lukas. El coeficiente de correlación de Pearson entre el instrumento diseñado y el PIL fue de 0.541, valor que se considera moderado. El coeficiente de correlación de Spearman's ρ fue igual a 0.42 entre el Logo Test y su instrumento, lo cual indica dos cuestiones: los tres instrumentos no miden la misma variable, pero sí conceptos relacionados.

La evidencia de la validez de constructo se obtiene mediante el análisis de factores. Tal método nos indica cuántas dimensiones integran a una variable y qué ítems conforman cada dimensión. Los reactivos que no pertenezcan a una dimensión, quiere decir que están "aislados" y no miden lo mismo que los demás ítems, por lo tanto deben eliminarse. Es un método que tradicionalmente se ha considerado complejo, por los cálculos estadísticos implicados, pero que es relativamente sencillo de interpretar y como los cálculos hoy en día los realiza la computadora, está al alcance de cualquier persona que se inicie dentro de la investigación. Este método se revisa —con ejemplos reales— en el capítulo 8 del CD anexo: Análisis estadístico-multivariado de los datos.

Análisis de factores Método estadístico multivariado que sirve para determinar el número y la naturaleza de un grupo de *constructos* subyacentes en un conjunto de mediciones.

La confiabilidad se obtiene en Minitab siguiendo los comandos: Stat ▶ Reliability/Survival, y en SPSS a través de: Analizar → Escalas → Análisis de fiabilidad. Las correlaciones que se desee calcular, depende del nivel de medición de las variables, pero ambos programas tienen varias opciones en Stat (Minitab) y Analizar (SPSS). El análisis de factores en Minitab se localiza en Stat → Multivariate, y en SPSS en Analizar → Reducción de datos → Análisis factorial. En las futuras versiones de estos programas, las opciones podrían cambiar, pero es cuestión de localizar en dónde se solicita el análisis de interés.

Una vez que se determina la confiabilidad (de 0 a 1) y se muestra la evidencia sobre la validez, si algunos ítems son problemáticos (no discriminan, no se vinculan a otros ítems, van en sentido contrario a toda la escala, no miden lo mismo, etc.), se eliminan de los cálculos (pero en el reporte de la investigación, se indica cuáles fueron eliminados, las razones de ello y cómo alteran los resultados); posteriormente se vuelve a realizar el análisis descriptivo (distribución de frecuencias, medidas de tendencia central y de variabilidad, etcétera).

En el CD anexo se presenta completa la validación de la escala de satisfacción sexual de Álvarez Gayou *et al.* (2005). Un estudio muy interesante y que muestra todos los elementos para generar un instrumento de medición, paso por paso. Incluye la generación de redes semánticas. Su abordaje es desde el punto de vista de la salud y con propiedad científica. Se prefirió incluirlo en el CD para presentar el trabajo completo.

¿Hasta aquí llegamos?

Cuando el estudio tiene una finalidad puramente exploratoria o descriptiva, debemos interrogarnos: ¿podemos establecer relaciones entre variables? En caso de una respuesta positiva, es factible seguir; pero si dudamos o el alcance se limitó a explorar y describir, el trabajo de análisis concluye y debemos comenzar a preparar el reporte de la investigación. De lo contrario es necesario continuar con la estadística inferencial.

Paso 5: analizar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial)

Estadística inferencial: de la muestra a la población

¿Para qué es útil la estadística inferencial?

Con frecuencia, el propósito de la investigación va más allá de describir las distribuciones de las variables: se pretende generalizar los resultados obtenidos en la muestra a la población o el universo. Los datos casi siempre se recolectan de una muestra y sus resultados estadísticos se denominan **estadígrafos**; la media o la desviación estándar de la distribución de una muestra son estadígrafos. A las estadísticas de la población o al universo se les conoce como **parámetros**. Los parámetros no son calculados, porque no se recolectan datos de toda la población, pero pueden ser inferidos de los estadígrafos, de ahí el nombre de **estadística inferencial**. El procedimiento de esta naturaleza de la estadística se esquematiza en la figura 10.16.

Estadística inferencial Se utiliza para probar hipótesis y estimar parámetros.

La inferencia de los parámetros se lleva a cabo mediante técnicas estadísticas apropiadas. Estas técnicas se explicarán más adelante.

La estadística inferencial se utiliza para dos procedimientos (Wiersma y Jurs, 2005):

- a) *Probar hipótesis*
- b) *Estimar parámetros*

¿En qué consiste la prueba de hipótesis?

Una hipótesis en el contexto de la estadística inferencial es una proposición respecto a uno o varios parámetros, y lo que el investigador hace por medio de la **prueba de hipótesis** es determinar si la hipótesis es congruente con los datos obtenidos en la muestra (Wiersma y Jurs, 2005).

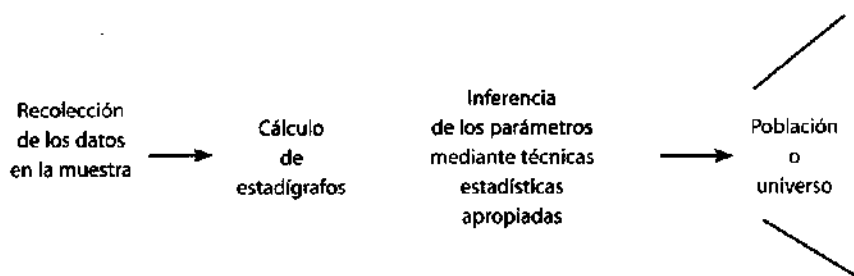


Figura 10.16 Procedimiento de la estadística inferencial.

La hipótesis se retiene como un valor aceptable del parámetro, si es congruente con los datos. Si no lo es, se rechaza (pero los datos no se descartan). Para comprender lo que es la prueba de hipótesis en la estadística inferencial es necesario revisar el concepto de distribución muestral⁴ y nivel de significancia.

¿Qué es una distribución muestral?

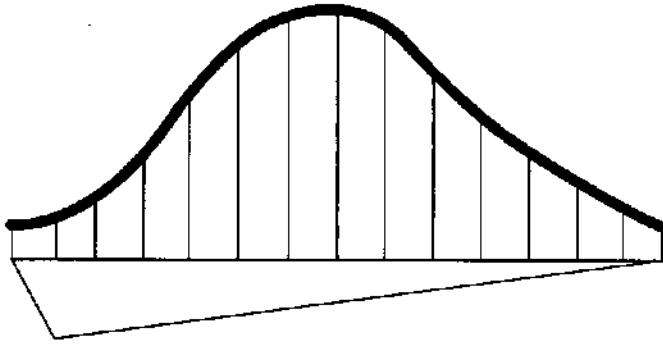
Una **distribución muestral** es un conjunto de valores sobre una estadística calculada de todas las muestras posibles de determinado tamaño de una población (Wiersma y Jurs, 2005). Las distribuciones muestrales de medias son probablemente las más conocidas. Explicemos este concepto con un ejemplo. Supongamos que nuestro universo son los automovilistas de una ciudad y deseamos averiguar cuánto tiempo pasan diariamente manejando ("al volante"). De este universo podría extraerse una muestra representativa. Vamos a suponer que el tamaño adecuado de muestra es de 512 automovilistas ($n = 512$). Del mismo universo se podrían extraer diferentes muestras, cada una con 512 personas.

Teóricamente, incluso podría elegirse al azar una, dos, tres, cuatro muestras, y las veces que fuera necesario hacerlo, hasta agotar todas las muestras posibles de 512 automovilistas de esa ciudad (todos los sujetos serían seleccionados en varias muestras). En cada muestra se obtendría una media del tiempo que pasan los automovilistas manejando. Tendríamos pues, una gran cantidad de medias, tantas como las muestras extraídas ($\bar{X}_1, \bar{X}_2, \bar{X}_3, \bar{X}_4, \bar{X}_5, \dots, \bar{X}_k$). Y con éstas elaboraríamos una distribución de medias. Habría muestras que, en promedio, pasaran más tiempo "al volante" que otras. Este concepto se representa en la figura 10.17.

Si calculáramos la media de todas las medias de las muestras, obtendríamos el valor de la media poblacional.

Muy rara vez se obtiene la distribución muestral (la distribución de las medias de todas las muestras posibles). Es más bien un concepto teórico definido por la estadística para los investigadores. Lo que comúnmente hacemos es extraer una sola muestra.

⁴ Distribución muestral y distribución de una muestra son conceptos diferentes, esta última es resultado de los datos de nuestra investigación y es por variable.



Son medias ($\bar{X}$), no se trata de puntuaciones. Cada media representaría una muestra.

Figura 10.17 Distribución muestral de medias.

En el ejemplo de los automovilistas, sólo una de las líneas verticales de la distribución muestral presentada en la figura 10.17 es la media obtenida para nuestra única muestra seleccionada de 512 personas. Y la pregunta es: ¿nuestra media está cerca de la media de la distribución muestral?, debido a que si está cerca podremos tener una estimación precisa de la media poblacional (el parámetro poblacional es prácticamente el mismo que el de la distribución muestral). Esto se expresa en *el teorema central del límite*:

Si una población (no necesariamente normal) tiene de media m y de desviación estándar s , la distribución de las medias en el muestreo aleatorio realizado en esta población tiende, al aumentar n , a una distribución normal de media m y desviación estándar $\frac{s}{\sqrt{n}}$ donde n es el tamaño de muestra.

El teorema especifica que la **distribución muestral** tiene una media igual a la de la población, una varianza igual a la varianza de la población dividida entre el tamaño de muestra (su desviación estándar es $\sigma/\sqrt{n}$ y se distribuye normalmente). La desviación estándar (s) es un parámetro normalmente desconocido, aunque es posible estimarlo por la desviación estándar de la muestra.

El concepto de *distribución normal* es importante otra vez y se ofrece una breve explicación en la figura 10.18.

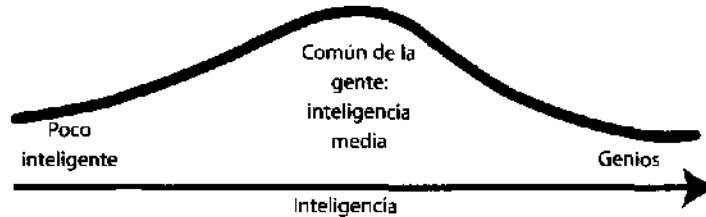
¿Qué es el nivel de significancia?

Wiersma y Jurs (2005) ofrecen una explicación sencilla del concepto, en la cual nos basaremos para analizar su significado.

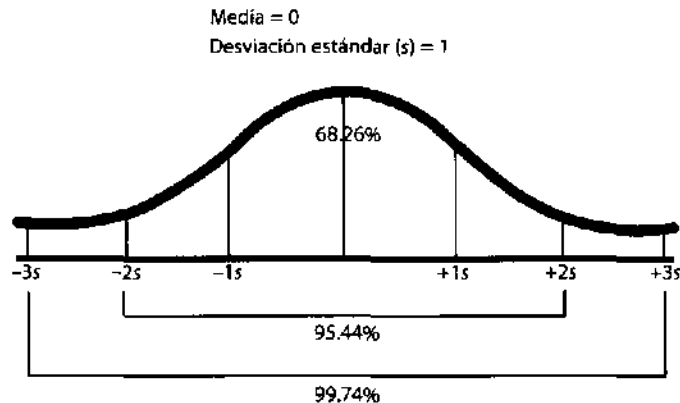
La probabilidad de que un evento ocurra oscila entre cero (0) y uno (1), donde cero significa la imposibilidad de ocurrencia y uno la certeza de que el fenómeno ocurra. Al lanzar al aire una moneda no cargada, la probabilidad de que salga “cruz” es de 0.50 y la probabilidad de que la moneda caiga en “cara” también es de 0.50. Con un dado, la probabilidad de obtener cualquiera de sus caras al lanzarlo es de $1/6 = 0.1667$. La suma de posibilidades siempre es de uno.

Nivel de significancia Es un nivel de la probabilidad de equivocarse y que fija de manera *a priori* el investigador.

Una gran cantidad de los fenómenos del comportamiento humano se manifiestan de la siguiente forma: la mayoría de las puntuaciones se concentran en el centro de la distribución, en tanto que en los extremos encontramos sólo algunas puntuaciones. Por ejemplo, la inteligencia: hay pocas personas muy inteligentes (genios), pero también hay pocas personas con muy baja inteligencia (por ejemplo: personas con capacidades mentales deficientes). La mayoría de los seres humanos somos medianamente inteligentes. Esto podría representarse así:



Debido a ello, se creó un modelo de probabilidad llamado curva normal o distribución normal. Como todo modelo es una distribución teórica que difícilmente se presenta en la realidad tal cual, pero sí se presentan aproximaciones a éste. La curva normal tiene la siguiente configuración:



68.26% del área de la curva normal es cubierta entre $-1s$ y $+1s$, 95.44% del área de esta curva es cubierta entre $-2s$ y $+2s$ y 99.74% se cubre con $-3s$ y $+3s$.

Las principales características de la distribución normal son:

1. Es *unimodal*, una sola moda.
2. La *asimetría es cero*. La mitad de la curva es exactamente igual a la otra mitad.
La distancia entre la media y $+3s$ es la misma que la distancia entre la media y $-3s$.
3. Es una función particular entre desviaciones con respecto a la media de una distribución y la probabilidad de que éstas ocurran.
4. La base está dada en *unidades de desviación estándar* (puntuaciones z), destacando las puntuaciones $-1s$, $-2s$, $-3s$, $+1s$, $+2s$ y $+3s$ (que equivalen respectivamente a $-1.00z$, $-2.00z$, $-3.00z$, $+1.00z$, $+2.00z$, $+3.00z$). Las distancias entre puntuaciones z representan áreas bajo la curva. De hecho, la distribución de puntuaciones z es la curva normal.
5. Es *mesocúrtica* (curtosis de cero).
6. La *media*, la *mediana* y la *moda* coinciden en el mismo punto.

Figura 10.18 Concepto de curva o distribución normal.

Aplicando el concepto de probabilidad a la distribución muestral, tomaremos el área de ésta como 1.00; en consecuencia, cualquier área comprendida entre dos puntos de la distribución corresponderá a la probabilidad de la distribución. Para probar hipótesis inferenciales respecto a la media, el investigador debe evaluar si es alta o baja la probabilidad de que la media de la muestra esté cerca de la media de la distribución muestral. Si es baja, el investigador dudará de generalizar a la población. Si es alta, el investigador podrá hacer generalizaciones. Es aquí donde entra el **nivel de significancia** o **nivel alfa** (α),⁵ el cual es un nivel de la probabilidad de equivocarse y se fija antes de probar hipótesis inferenciales.

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, ¿sentiría 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. Obtiene una estadística en una muestra (por ejemplo, la media) y analiza qué porcentaje tiene de confianza en que dicha estadística se acerque al valor de la distribución muestral (que es el valor de la población o el parámetro). Busca un alto porcentaje de confianza, una probabilidad elevada para estar tranquilo, porque sabe que tal vez haya error de muestreo y, aunque la evidencia parece mostrar una aparente “cercanía” entre el valor calculado en la muestra y el parámetro, tal “cercanía” puede no ser real o deberse a errores en la selección de la muestra.

¿Con qué porcentaje de confianza el investigador generaliza, para suponer que tal cercanía es real y no por un error de muestreo? *Existen dos niveles convenidos en ciencias sociales:*

- a) *El nivel de significancia de 0.05*, el cual implica que el investigador tiene 95% de seguridad para generalizar sin equivocarse y sólo 5% en contra. En términos de probabilidad, 0.95 y 0.05, respectivamente; ambos suman la unidad.
- b) *El nivel de significancia de 0.01*, el cual implica que el investigador tiene 99% en su favor y 1% en contra ($0.99 + 0.01 = 1.00$) para generalizar sin temor.

A veces el nivel de significancia puede ser todavía más riguroso, como al generalizar resultados de medicamentos o vacunas; o la resistencia de los materiales de un edificio (por ejemplo, 0.001, 0.00001, 0.00000001), pero al menos debe ser de 0.05. No se acepta un nivel de 0.06 (94% a favor de la generalización confiable), porque se busca hacer ciencia, no intuición.

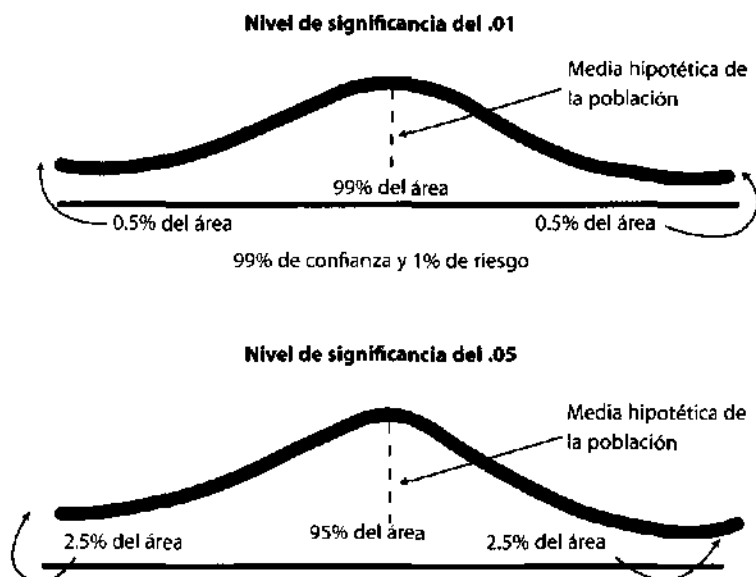
El *nivel de significancia* es un valor de certeza que el investigador fija *a priori*, respecto a no equivocarse. Cuando uno lee en un reporte de investigación que los resultados fueron significativos al nivel del 0.05 ($p < 0.05$), indica lo que se comentó: que existe 5% de posibilidad de error al aceptar la hipótesis, correlación o valor obtenido al aplicar una prueba estadística; o 5% de riesgo de que se rechace una hipótesis nula cuando era verdadera (Mertens, 2005).

Volveremos más adelante sobre este punto.

⁵ No confundir con el coeficiente Alfa-Cronbach, éste es para determinar la confiabilidad.

¿Cómo se relacionan la distribución muestral y el nivel de significancia?

El **nivel de significancia** se expresa en términos de probabilidad (0.05 y 0.01) y la distribución muestral también como probabilidad (el área total de ésta como 1.00). Pues bien, para ver si existe o no confianza al generalizar acudimos a la distribución muestral, con una probabilidad adecuada para la investigación. El nivel de significancia lo tomamos como un área bajo la distribución muestral, como se observa en la figura 10.19, y depende de si elegimos un nivel de 0.05 o de 0.01.



Nota:

1. Podemos expresarlo en proporciones (0.025, 0.95 y 0.025, respectivamente) o porcentajes como está en la gráfica.
2. 95% representa el área de confianza y 2.5%, el área de riesgo ($2.5\% + 2.5\% = 5\%$) en cada extremo, porque en nuestra estimación de la media poblacional pasaríamos hacia valores más altos o bajos.

Figura 10.19 Niveles de significancia en la distribución muestral.

Así, el nivel de significancia representa áreas de riesgo o confianza en la distribución muestral.

Una vez definido el nivel de significancia, ¿qué hacemos para ver si nuestra hipótesis sobre la media poblacional es aceptada o rechazada?

Antes de estudiar el procedimiento, es necesario hacer las siguientes consideraciones:

- a) La distribución muestral es una distribución normal de puntuaciones z , la base de la curva son puntuaciones z o unidades de desviación estándar.
- b) Las puntuaciones z son distancias que indican áreas bajo la distribución normal. En este caso, áreas de probabilidad.
- c) El área de riesgo es tomada como el área de rechazo de la hipótesis; por el contrario, el área de confianza, como el área de aceptación de la hipótesis.
- d) Se habla de una hipótesis acerca del parámetro (en este caso, media poblacional).

Si partimos de estas consideraciones, el procedimiento es:

1. Sobre bases firmes (revisión de la literatura e información disponible), establecer una hipótesis acerca del parámetro poblacional. Por ejemplo: El promedio de horas diarias que se exponen los niños de la ciudad de Valladolid a la televisión en fin de semana es de 3.0 horas.
2. Definir el nivel de significancia. Por ejemplo, $\alpha = 0.05$.
3. Recolectar los datos en una muestra representativa. Vamos a suponer que obtuvimos una media de 2.9 horas y una desviación estándar de 1.2 horas; la muestra incluyó 312 niños.
4. Estimar la desviación estándar de la distribución muestral de la media utilizando la siguiente fórmula:

$$S_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Donde $S_{\bar{x}}$ es la desviación estándar de la distribución muestral de la media, s representa la desviación estándar de la muestra y n es el tamaño de la muestra.

En el ejemplo:

$$S_{\bar{x}} = \frac{1.2}{\sqrt{312}}$$

$$S_{\bar{x}} = 0.0679$$

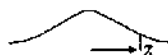
5. Transformar la media de la muestra en una puntuación z , en el contexto de la distribución muestral, con una variación de la fórmula ya conocida para obtener puntuaciones z :

$$z = \frac{X - \bar{X}}{S_{\bar{x}}}$$

donde X es la media de la muestra (recordemos que la distribución muestral es de medias y no de puntuaciones), $\bar{X}$ es la media hipotética de la distribución muestral (parámetro poblacional), $S_{\bar{x}}$ es la desviación estándar de la distribución muestral de medias. Así, tenemos:

$$z = \frac{2.9 - 3.0}{0.0679} = -1.47$$

6. En la *tabla de áreas bajo la curva normal* (ver apéndice 4, tabla 1 del CD),⁶ buscar aquella puntuación z que deje a 2.5% por encima de ella, que es 1.96. En la tabla 1 del apéndice 4 se presenta la distribución de puntuaciones z , sólo la mitad, pues debemos recordar que es una distribución simétrica y se aplica igual para ambos lados de la media. Así se incluye en los textos de estadística. Se busca 2.5%, porque la tabla sólo abarca la mitad de la distribución y el riesgo que estamos afrontando es de 5% (2.5% del extremo de cada lado). La tabla mencionada contiene cuatro columnas: la primera indica puntuaciones z , la segunda expresa la distancia de la puntuación z a la media. La tercera, el área que está por debajo de esa puntuación desde el comienzo de la distribución, como se muestra en la gráfica:



Y la cuarta, el área que está por encima de esa puntuación:



Las áreas se expresan en proporciones. Lo que buscamos es una puntuación z que deje por encima un área de 0.0250 o 2.5% (la encontramos en la cuarta columna de la tabla); esta puntuación z es 1.96. Siempre que nuestro nivel de significancia sea 0.05, tomamos la puntuación z de 1.96.

7. Comparar la media de la muestra transformada a puntuación z con el valor 1.96; si es menor, aceptar la hipótesis; si es mayor, rechazarla. Veamos el ejemplo:

$$\frac{\text{Media de la muestra transformada a } z}{1.47}$$

$$\frac{\text{Nivel de significancia del 0.05}}{\pm 1.96}$$

Decisión: Aceptar la hipótesis a un nivel de significancia de 0.05 (95% a favor y 5% de riesgo de cometer un error).

Si la media obtenida,
al transformarse en z ,
hubiera sido: 3.25,

7.46 o un valor mayor → Rechazar la hipótesis

Por ejemplo:

Media de la muestra = 2.0

Desviación estándar de la muestra = 0.65

$$n = 700$$

$$S\bar{x} = 0.0246$$

$$z = 40.65$$

La media está situada a más de 40 desviaciones estándar de la media; se localiza en la zona crítica (más allá de 1.96 desviaciones estándar). Rechazar la hipótesis.

⁶ Material incluido en el CD anexo. SPSS trabaja áreas bajo la curva con los comandos: Gráficos → Curva COR.

¿Por qué es importante otro concepto: el intervalo de confianza?

Se ha hablado de la distribución muestral por lo que respecta a la prueba de hipótesis, pero otro procedimiento de la estadística inferencial es construir un *intervalo* donde se localiza un parámetro (Wiersma y Jurs, 2005). Por ejemplo, en lugar de pretender probar una hipótesis acerca de la media poblacional, puede buscarse un intervalo donde se ubique dicha media. Lo anterior requiere un nivel de confianza, al igual que en la prueba de hipótesis inferenciales. El nivel de confianza es al intervalo de confianza, lo que el nivel de significancia es a la prueba de hipótesis. Es decir, se trata de una probabilidad definida de que un parámetro se va a ubicar en un determinado intervalo. Los niveles de confianza más comunes en la investigación son 0.95 y 0.99. Su sentido es similar al ya comentado, si es de 0.95; quiere decir que tenemos 95% en favor de que el parámetro se localice en el intervalo estimado, contra 5% de elegir un intervalo equivocado. El nivel de 0.99 señala 99% de probabilidad de seleccionar el intervalo adecuado. Tales niveles de confianza se expresan en unidades de desviación estándar. Una vez más se acude a la distribución muestral, concretamente a la tabla de áreas bajo la curva normal (apéndice 4, tabla 1 del CD anexo), y se elige la puntuación z correspondiente al nivel de confianza seleccionado. Una vez hecho esto, se aplica la siguiente fórmula:

$$\text{Intervalo de confianza} = \text{estadígrafo} + \left(\begin{array}{c} \text{Puntuación } z \text{ que} \\ \text{expresa el nivel de} \\ \text{confianza elegido} \end{array} \right) \left(\begin{array}{c} \text{Desviación estándar de la} \\ \text{distribución muestral} \\ \text{correspondiente} \end{array} \right)$$

En la fórmula, el estadígrafo es la estadística calculada en la muestra, la puntuación z es 1.96 con un nivel de 0.95 y de 2.58 con un nivel de 0.99, en tanto que el error estándar depende del estadígrafo en cuestión. Veámoslo con el ejemplo de la media en el caso de la exposición diaria a la televisión (en fin de semana) por parte de los niños de Valladolid:

Media = 2.9 horas

$s = 1.2$ horas

$\bar{Sx} = 0.0679$

(desviación estándar
de la distribución
muestral de la media).

Nivel de confianza = 0.95 ($z = 1.96$)

Intervalo de confianza = $2.9 \pm (1.96) (0.0679)$
= $2.9 \pm (0.133)$

Intervalo de confianza: la media poblacional está entre 2.767 y 3.033 horas, con 95% de probabilidades de no cometer error.

¿Se pueden cometer errores al realizar estadística inferencial?

Nunca estaremos completamente seguros de nuestra estimación. Trabajamos con altos niveles de confianza o seguridad, pero, aunque el riesgo es mínimo, podría cometerse un error. *Los resultados posibles al probar hipótesis serían:*

1. Aceptar una hipótesis verdadera (decisión *correcta*).
2. Rechazar una hipótesis falsa (decisión *correcta*).
3. Aceptar una hipótesis falsa (conocido como *error del Tipo II o error beta*).
4. Rechazar una hipótesis verdadera (conocido como *error del Tipo I o error alfa*).

Ambos tipos de error son indeseables; sin embargo, puede *reducirse la posibilidad* de que se presenten mediante:

- a) *Muestras representativas probabilísticas.*
- b) *Inspección cuidadosa de los datos.*
- c) *Selección de las pruebas estadísticas apropiadas.*
- d) *Mayor conocimiento de la población.*

*** Prueba de hipótesis

Hay dos tipos de análisis estadísticos que pueden realizarse para probar hipótesis: los *análisis paramétricos* y los *no paramétricos*. Cada tipo posee sus características y presuposiciones que lo sustentan; la elección de qué clase de análisis efectuar depende de estas presuposiciones. De igual forma, cabe destacar que en una misma investigación es posible llevar a cabo análisis paramétricos para algunas hipótesis y variables y análisis no paramétricos para otras. Asimismo, los análisis a realizar dependen de las hipótesis que hayamos formulado y el nivel de medición de las variables que las conforman. Primero revisaremos las pruebas más importantes y luego la secuencia de análisis según las hipótesis establecidas.

*** Análisis paramétricos

¿Cuáles son los supuestos o las presuposiciones de la estadística paramétrica?

Para realizar **análisis paramétricos** debe partirse de los siguientes supuestos:

1. *La distribución poblacional de la variable dependiente es normal:* el universo tiene una distribución normal.
2. *El nivel de medición de la variable dependiente es por intervalos o razón.*
3. *Cuando dos o más poblaciones son estudiadas, tienen una varianza homogénea:* las poblaciones en cuestión poseen una dispersión similar en sus distribuciones (Wiersma y Jurs, 2005).

Ciertamente estos criterios son tal vez demasiado rigurosos y algunos investigadores sólo basan sus análisis en el tipo de hipótesis y los niveles de medición de las variables. Esto queda a juicio del lector. En la investigación académica y cuando quien la realiza es una persona experimentada, sí debe solicitársele tal rigor.

¿Cuáles son los métodos o las pruebas estadísticas paramétricas más utilizadas?

Las pruebas estadísticas paramétricas más utilizadas son:

- Coeficiente de correlación de Pearson y regresión lineal
- Prueba t
- Prueba de contraste de la diferencia de proporciones
- Análisis de varianza unidireccional (ANOVA *Oneway*)
- Análisis de varianza factorial (ANOVA)
- Análisis de covarianza (ANOVA)

¿Qué es el coeficiente de correlación de Pearson?

Definición: Es una prueba estadística para analizar la relación entre dos variables medidas en un nivel por intervalos o de razón.

Se simboliza: r

Hipótesis a probar: Correlacional, del tipo de “A mayor X, mayor Y”, “A mayor X, menor Y”, “altos valores en X están asociados con altos valores en Y”, “altos valores en X se asocian con bajos valores de Y”.

Variables: Dos. La prueba en sí no considera a una como independiente y a otra como dependiente, ya que no evalúa la causalidad. La noción de causa-efecto (independiente-dependiente) es posible establecerla teóricamente, pero la prueba no considera dicha causalidad.

El coeficiente de correlación de Pearson se calcula a partir de las puntuaciones obtenidas en una muestra en dos variables. Se relacionan las puntuaciones obtenidas de una variable con las puntuaciones obtenidas de la otra, con los mismos participantes o casos.

Nivel de medición de las variables: Intervalos o razón.

Interpretación: El coeficiente r de Pearson puede variar de -1.00 a $+1.00$, donde:

- -1.00 = *correlación negativa perfecta*. (“A mayor X, menor Y”, de manera proporcional. Es decir, cada vez que X aumenta una unidad, Y disminuye siempre una cantidad constante.) Esto también se aplica “a menor X, mayor Y”.
- -0.90 = Correlación negativa muy fuerte.
- -0.75 = Correlación negativa considerable.
- -0.50 = Correlación negativa media.
- -0.25 = Correlación negativa débil.
- -0.10 = Correlación negativa muy débil.
- 0.00 = No existe correlación alguna entre las variables.
- $+0.10$ = Correlación positiva muy débil.
- $+0.25$ = Correlación positiva débil.
- $+0.50$ = Correlación positiva media.
- $+0.75$ = Correlación positiva considerable.
- $+0.90$ = Correlación positiva muy fuerte.
- $+1.00$ = *Correlación positiva perfecta*.

"A mayor X , mayor Y " o "a menor X , menor Y ", de manera proporcional. (Cada vez que X aumenta, Y aumenta siempre una cantidad constante.)

El signo indica la dirección de la correlación (positiva o negativa); y el valor numérico, la magnitud de la correlación. Los principales programas computacionales de análisis estadístico reportan si el coeficiente es o no significativo de la siguiente manera:

$s = 0.001$	significancia
0.7831	valor del coeficiente

Si s es menor del valor 0.05, se dice que el coeficiente es *significativo* en el nivel de 0.05 (95% de confianza en que la correlación sea verdadera y 5% de probabilidad de error). Si s es menor a 0.01, el coeficiente es *significativo* al nivel de 0.01 (99% de confianza de que la correlación sea verdadera y 1% de probabilidad de error).

O bien otros programas como SPSS los presentan en una tabla, se señala con asterisco el nivel de significancia: donde un asterisco (*) implica una significancia menor a 0.05 (quiere decir que el coeficiente es significativo en el nivel de 0.05, la probabilidad de error es menor de 5%) y dos asteriscos (**) una significancia menor a 0.01 (la probabilidad de error es menor de 1%).

Podemos verlo en el ejemplo de la tabla 10.12:

Tabla 10.12 Correlaciones entre moral y dirección

		Correlaciones	
		Moral	Dirección
Moral	Correlación de Pearson	1	0.557**
	Sig. (bilateral)		0.000
	N	362	335
Dirección	Correlación de Pearson	0.557**	1
	Sig. (bilateral)	0.000	
	N	335	373

** La correlación es significativa al nivel 0.01 (bilateral).

Obsérvese que se correlacionan dos variables: "moral" y "dirección", aunque la correlación aparece dos veces, porque es una tabla que hace todas las comparaciones posibles entre las variables y al hacerlo, genera un eje diagonal (representado por las correlaciones de las variables contra ellas mismas —"moral" con "moral", y "dirección" con "dirección", que carece de sentido porque son las mismas puntuaciones, por eso es perfecta—), y por encima de ese eje aparecen todos los coeficientes, y se repiten por debajo del eje. La correlación es de 0.557 y es significativa en el nivel del 0.000 (menor del 0.01). N representa el número de casos correlacionados.

Una correlación de Pearson puede ser significativa, pero si es menor a 0.30 no resulta tan útil.

Consideraciones: Cuando el coeficiente r de Pearson se eleva al cuadrado (r^2), se obtiene el coeficiente de determinación y el resultado indica la *varianza de factores comunes*. Esto es, el porcentaje de la variación de una variable debido a la variación de la otra variable y viceversa (o cuánto explica o determina una variable la variación de la otra). Veámoslo gráficamente en la figura 10.20.

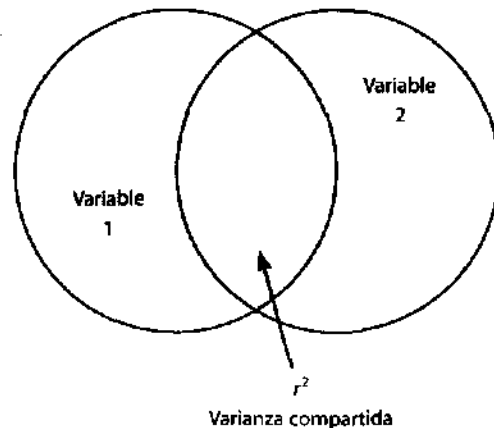


Figura 10.20 Varianza de factores comunes.

Por ejemplo: Si la correlación entre “productividad” y “asistencia al trabajo” es de 0.80.

$$r = 0.80$$

$$r^2 = 0.64$$

“La productividad” constituye a, o explica, 64% de la variación de “la asistencia al trabajo”. “La asistencia al trabajo” explica 64% de “la productividad”. Si r es 0.72 y consecuentemente $r^2 = 0.52$, quiere decir que poco más de la mitad de la variabilidad en una variable está explicada por la otra.

EJEMPLO

Hi:	A mayor motivación intrínseca, mayor puntualidad.
Resultado:	$r = 0.721$ $s = 0.0001$
Interpretación:	Se acepta la hipótesis de investigación en el nivel de 0.01. La correlación entre la motivación intrínseca y la productividad es considerable.
Hi:	A mayor ingreso, mayor motivación intrínseca.
Resultado:	$r = 0.214$ $s = 0.081$
Interpretación:	Se acepta la hipótesis nula. El coeficiente no es significativo: 0.081 es mayor que 0.05; recordemos que 0.05 es el nivel mínimo para aceptar la hipótesis.

Nota precatoria: Recuérdese lo referente a correlaciones espurias que se comentaron en el capítulo de tipos de estudio.

Creswell (2005) señala que un coeficiente de determinación (r^2) entre 0.66 y 0.85 ofrece una buena predicción de una variable respecto de la otra variable; y por encima de 0.85 implica que ambas variables miden casi el mismo concepto subyacente, son “cercanamente” la misma variable.

El coeficiente de correlación de Pearson es útil para relaciones lineales, como lo veremos en la regresión lineal, pero no para relaciones curvilíneas, en este caso, se suele usar Spearman rho (r_s).

Cuando queremos correlacionar simultáneamente más de dos variables (ejemplo: motivación, satisfacción en el trabajo, moral y autonomía), se utiliza el coeficiente de correlación múltiple o R , el cual se revisa en el capítulo 8 del CD anexo.

¿Qué es la regresión lineal?

Definición: Es un modelo matemático para estimar el efecto de una variable sobre otra. Está asociado con el coeficiente r de Pearson. Brinda la oportunidad de predecir las puntuaciones de una variable tomando las puntuaciones de la otra variable. Entre mayor sea la correlación entre las variables (covariación), mayor capacidad de predicción.

Hipótesis: Correlacionales y causales.

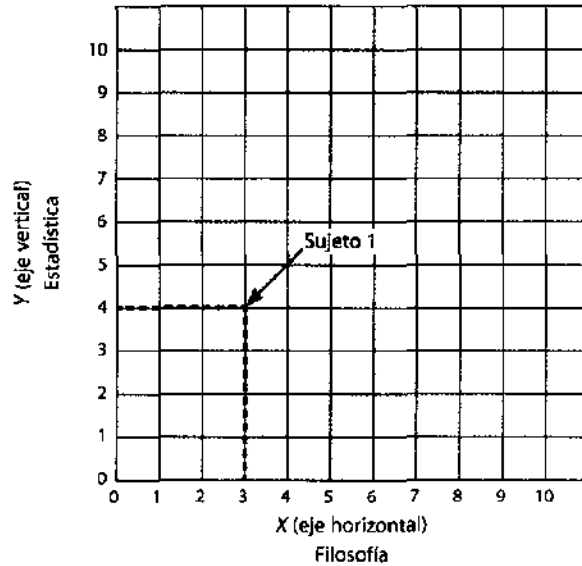
Variables: Dos. Una se considera como independiente y otra como dependiente. Pero, para poder hacerlo, debe tenerse un sólido sustento teórico.

Nivel de medición de las variables: Intervalos o razón.

Procedimiento e interpretación: La regresión lineal se determina con base en el diagrama de dispersión. Éste consiste en una gráfica donde se relacionan las puntuaciones de una muestra en dos variables. Veámoslo con un ejemplo sencillo de ocho casos. Una variable es la calificación en Filosofía y la otra variable es la calificación en Estadística; ambas medidas, hipotéticamente, de 0 a 10.

<i>Sujetos</i>	<i>Puntuaciones Filosofía (X)</i>	<i>Estadística (Y)</i>
1	3	4
2	8	8
3	9	8
4	6	5
5	10	10
6	7	8
7	6	7
8	5	5

El *diagrama de dispersión* se construye graficando cada par de puntuaciones en un espacio o plano bidimensional. Sujeto "1" tuvo 3 en X (filosofía) y 4 en Y (estadística):



Así se grafican todos los pares:

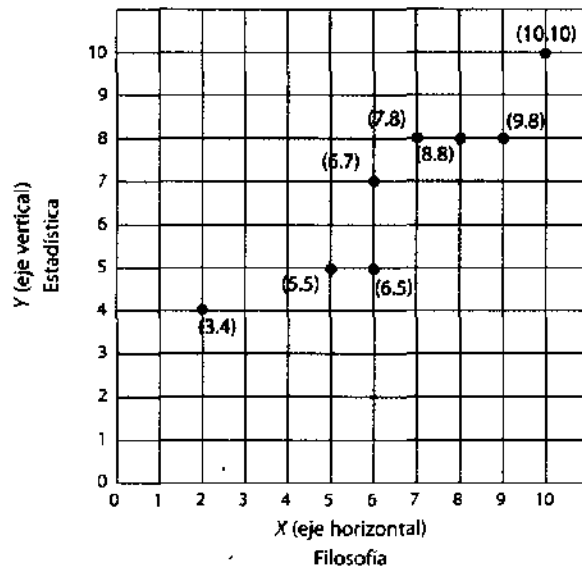
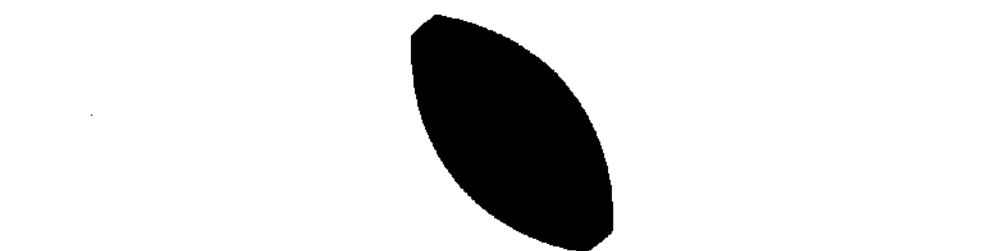


Figura 10.21 Ejemplos de gráficas de dispersión.

Los *diagramas de dispersión* son una manera de visualizar gráficamente una correlación. Por ejemplo:



Correlación positiva muy fuerte: la tendencia es ascendente, altas puntuaciones en *X* y altas puntuaciones en *Y*.

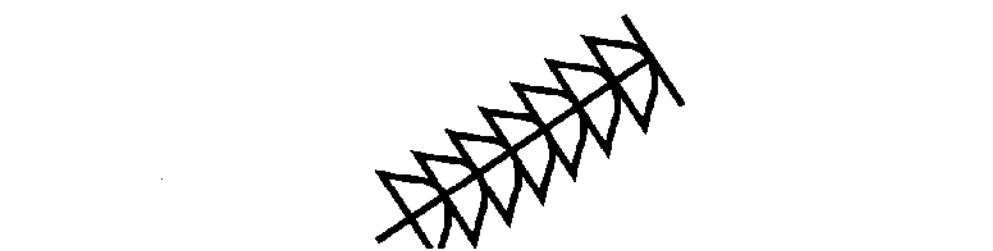


Correlación negativa considerable



Ausencia de correlación

Así, cada punto representa un caso y un resultado de la intersección de las puntuaciones en ambas variables. El diagrama de dispersión puede ser resumido a una línea.



Conociendo la línea y la tendencia, podemos predecir los valores de una variable conociendo los de la otra variable.

Figura 10.21 Ejemplos de gráficas de dispersión (continuación).

Esta línea se expresa mediante la *ecuación de regresión lineal*:

$$Y = a + bX$$

en donde Y es un *valor de la variable dependiente* que se desea predecir, a es la *ordenada* en el origen y b la *pendiente* o inclinación, X es el valor que fijamos en la variable independiente.

Los programas y paquetes computacionales de análisis estadístico que incluyen la *regresión lineal*, proporcionan los datos de a y b .

a o *intercept* y b o *slope*

Para predecir un valor de Y , se sustituyen los valores correspondientes en la ecuación.

EJEMPLO

$$a \text{ (intercept)} = 1.2$$

$$b \text{ (slope)} = 0.8$$

Entonces podemos hacer la predicción: ¿a un valor de 7 en Filosofía qué valor le corresponde en Estadística?

$$Y = \frac{1.2}{a} + \frac{(0.8)}{b} \quad \frac{(7)}{X}$$

$$Y = 6.8$$

Predecimos que a un valor de 7 en X le corresponderá un valor de 6.8 en Y .

EJEMPLO

Regresión lineal

Hi: La autonomía laboral es una variable para predecir la motivación intrínseca en el trabajo. Ambas variables están relacionadas.

Las dos variables fueron medidas en una escala por intervalos de 1 a 5.

Resultado: $a \text{ (intercept)} = 0.42$

$$b \text{ (slope)} = 0.65$$

Interpretación: Cuando X (autonomía) es 1, la predicción estimada de Y es 1.07; cuando X es 2, la predicción estimada de Y es 1.72; cuando X es 3, Y será 2.37; cuando X es 4, Y será 3.02; y cuando X es 5, Y será 3.67.

$$Y = a + bX$$

$$1.07 = 0.42 + 0.65 \text{ (1)}$$

$$1.72 = 0.42 + 0.65 \text{ (2)}$$

$$2.37 = 0.42 + 0.65 \text{ (3)}$$

$$3.02 = 0.42 + 0.65 \text{ (4)}$$

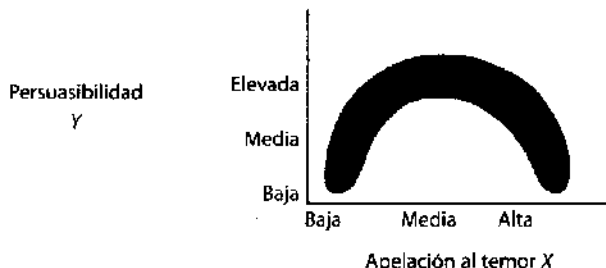
$$3.67 = 0.42 + 0.65 \text{ (5)}$$

Consideraciones: La *regresión lineal* es útil con relaciones lineales, no con *relaciones curvilineales*. Porque como señalan León y Montero (2003, p.191) es un error atribuir a la relación causal una covariación exclusivamente lineal: a mayores valores en la variable independiente, mayores valores en la dependiente. Existen muchas relaciones de causa-efecto que no son lineales, como por ejemplo: la vinculación entre ansiedad y rendimiento. Cierta grado de ansiedad ayuda a conseguir mejores resultados en un examen o la práctica de un deporte; pero, por encima de cierto nivel (nerviosismo extremo), la ejecución empeora. En la figura 10.22 se muestran ejemplos de estas relaciones.

Las *relaciones curvilineales* son aquellas en las cuales la tendencia varía: primero es ascendente y luego descendente, o viceversa.

Se ha demostrado que una estrategia persuasiva con niveles altos de apelación al temor, por ejemplo, un comercial televisivo muy dramático, provoca una baja persuasibilidad, lo mismo que una estrategia persuasiva con niveles muy bajos de apelación al temor.

La estrategia persuasiva más adecuada es la que utiliza niveles medios de apelación al temor. Esta relación es curvilineal; se representaría así:



Otras gráficas de relaciones curvilineales serían:

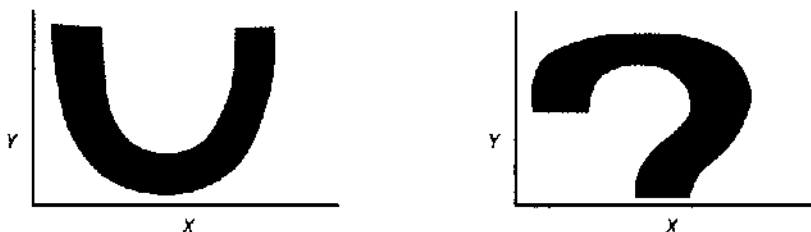


Figura 10.22 Ejemplos de relaciones curvilineales.

¿Qué es la prueba *t*?

Definición: Es una prueba estadística para evaluar si dos grupos difieren entre sí de manera significativa respecto a sus medias.

Se simboliza: *t*

Hipótesis: De diferencia entre dos grupos. La hipótesis de investigación propone que los grupos difieren de manera significativa entre sí y la hipótesis nula propone que los grupos no difieren significativamente. Los grupos pueden ser dos plantas comparadas en su productividad, dos es-

cuelas contrastadas en los resultados a un examen, dos clases de materiales de construcción comparados en su rendimiento, etcétera.

Variable: La comparación se realiza sobre una variable (teóricamente dependiente). Si hay diferentes variables, se efectuarán varias pruebas t (una por cada variable), y la razón que motiva la creación de los grupos puede ser una variable independiente. Por ejemplo, un experimento con dos grupos, donde a uno se le aplica el estímulo experimental y al otro no, es de control.

Nivel de medición de la variable de comparación: Intervalos o razón.

Interpretación: El valor t se obtiene en muestras grandes mediante la fórmula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

En donde $\bar{X}_1$ es la media del primer grupo, $\bar{X}_2$ la media del segundo grupo, s_1^2 representa la desviación estándar del primero elevada al cuadrado, n_1 es el tamaño del primer grupo, s_2^2 simboliza la desviación estándar del segundo grupo elevada al cuadrado y n_2 es el tamaño del segundo grupo. En realidad, el denominador es el *error estándar de la distribución muestral de la diferencia entre medias*.

Para saber si el valor t es significativo, se aplica la fórmula y se calculan los grados de libertad. La prueba t se basa en una distribución muestral o poblacional de diferencia de medias conocida como la distribución t de Student. Esta distribución se identifica por los grados de libertad, los cuales constituyen el número de maneras en que los datos pueden variar libremente. Son determinantes, ya que nos indican qué valor debemos esperar de t , dependiendo del tamaño de los grupos que se comparan. *Cuanto mayor número de grados de libertad se tengan, la distribución t de Student se acercará más a ser una distribución normal* y usualmente, si los grados de libertad exceden los 120, la distribución normal se utiliza como una aproximación adecuada de la distribución t de Student (Wiersma y Jurs, 2005).

Los grados de libertad se calculan con la fórmula siguiente, en la que n_1 y n_2 son el tamaño de los grupos que se comparan:

$$gl = (n_1 + n_2) - 2$$

Vogt (1999) señala que los grados de libertad indican cuántos casos fueron usados para calcular un valor estadístico en particular.

Una vez calculados el valor t y los grados de libertad, se elige el *nivel de significancia* y se compara el valor obtenido contra el valor que le correspondería, en la tabla 2 del apéndice 4 del CD anexo (*tabla de la distribución t de Student*). Si el valor calculado es igual o mayor al que aparece en la tabla, se acepta la hipótesis de investigación. Pero si es menor, se acepta la hipótesis nula.

En la tabla se busca el valor con el cual vamos a comparar el que hemos calculado, con base en el nivel de confianza elegido (0.05 o 0.01) y los grados de libertad. La tabla contiene los niveles de confianza como columnas y los grados de libertad como renglones. Los niveles de confianza adquieren el significado del que se ha hablado (el 0.05 significa 95% de que los grupos en realidad difieran significativamente entre sí y 5% de posibilidad de error).

Cuanto mayor sea el valor t calculado respecto al valor de la tabla y menor sea la posibilidad de error, mayor será la certeza en los resultados.

Consideraciones: La prueba t se utiliza para comparar los resultados de una preprueba con los resultados de una posprueba en un contexto experimental. Se comparan las medias y las varianzas del grupo en dos momentos diferentes: $(\bar{X}_1) \times (\bar{X}_2)$. O bien, para comparar las prepruebas o pospruebas de dos grupos que participan en un experimento:

$$\begin{array}{lcl} G_1 & (\bar{X}_1) & \\ & & t \\ G_2 & (\bar{X}_2) & \text{O son las pospruebas} \end{array}$$

EJEMPLO

Algunos ejemplos de la aplicación de este método podrían ser las sesiones en donde se evalúe un nuevo producto, digamos un dulce. En estas sesiones se les proporciona el producto a los participantes y se les solicitan opiniones sobre sabor, color, presentación, precio, cualidades del producto; se discuten a fondo las propiedades, cualidades y carencias del producto; se administra una escala de actitudes o un cuestionario, y se hacen preguntas abiertas a cada participantes. O bien, sesiones para analizar la popularidad de varios candidatos políticos, evaluar el servicio y la atención recibida en un hospital, indagar la percepción de un grupo de estudiantes sobre la calidad de la enseñanza recibida o conocer la opinión de los sectores de una comunidad sobre una reforma electoral.

Otro caso sería llevar a cabo sesiones con parejas que no pudieron tener hijos biológicos y decidieron adoptar, con el propósito de explorar el significado de la paternidad-maternidad en ellos; indagar sobre sus razones profundas para tomar la decisión de adoptar; conocer los sentimientos y las emociones que experimentaron antes de la adopción, durante el proceso y después que concluyó éste; apreciar su estado actual de ánimo, su sentido de vida, su percepción de sí mismos y su relación de pareja, así como evaluar (no en sentido cuantitativo) la interacción con el (la) hijo(a) adoptado(a).

Y recordamos que se pueden organizar sesiones con diferentes tipos de población y mixtas. Por ejemplo, los autores participaron en un estudio para evaluar un programa televisivo que acababa de salir al aire con un nombre y formato nuevos. Además de realizarse una encuesta telefónica, se organizaron varias sesiones (algunas con amas de casa, otras con estudiantes, también con trabajadores, empleados de oficina, ejecutivos, profesores, publicistas y otros grupos tipo, así como reuniones donde participaban, por ejemplo, una ama de casa de más de 60 años, una ama de casa de 50 años, una ama de casa más joven, un empleado de una oficina pública, una secretaria, un profesor, un dependiente de supermercado, un ejecutivo y dos estudiantes.

En las sesiones se profundizó en el formato del programa (música, manejo de cámaras, duración, inserción de comerciales, sonido, manera de presentar invitados...), conductores, contenido, etcétera.

Al acudir a la tabla de la distribución t de *Student* (tabla 2, apéndice 4, incluido en el CD), buscamos los grados de libertad correspondientes y elegimos en la columna de gl , el renglón α , que se selecciona siempre que se tienen más de 200 grados de libertad. La tabla contiene los siguientes valores:

gl	0.05	0.01
α (mayor de 200)	1.645	2.326

Nuestro valor calculado de t es 6.698 y resulta superior al valor de la tabla en un nivel de confianza de 0.05 ($6.698 > 1.645$). Entonces, la conclusión es que aceptamos la hipótesis de investigación y rechazamos la nula. Incluso, el valor t calculado es superior en un nivel de confianza del 0.01 ($6.698 > 2.326$).

Comentario: Efectivamente, en el contexto de la investigación (del ejemplo anterior), los varones le atribuyen más importancia al atractivo físico en sus relaciones de pareja que las mujeres.

Si tuviéramos 60 grados de libertad y un valor t igual a 1.27, al comparar este valor con los de la tabla obtendríamos:

gl	0.05	0.01
60	1.6707	2.390

El valor t calculado es menor a los valores de la tabla. Se *rechaza la hipótesis de investigación y se acepta la hipótesis nula*.

Cuando el valor t se calcula mediante un paquete estadístico computacional, la significancia se proporciona como parte de los resultados y ésta debe ser menor a 0.05 o 0.01, lo cual depende del nivel de confianza seleccionado (en SPSS se ofrece el resultado en dos versiones, según sea el caso, si se asumen o no varianzas iguales). Lo más importante es visualizar el valor t y su significancia. Ver tabla 10.13.

El tamaño del efecto. Al comparar grupos en este caso con la prueba t es importante determinar el *tamaño del efecto*, que es una medida de la “fuerza” de la diferencia de las medias u otros valores considerados (Creswell, 2005). Resulta ser una medida en unidades de desviación estándar.

¿Cómo se calcula?

El tamaño del efecto es justo la diferencia estandarizada entre las medias de los dos grupos. En otras palabras:

$$\text{Tamaño del efecto} = \frac{\text{Media del grupo 1} - \text{Media del grupo 2}}{\text{Desviación estándar sopesada}}$$

La desviación estándar sopesada (Creswell, 2005) es la estimación reunida de la desviación estándar de ambos grupos, basada en la premisa que cualquier diferencia entre sus desviaciones es solamente debida a la variación del muestreo.

La desviación estándar sopesada (denominador en la fórmula) se calcula así:

$$\sqrt{\frac{(N_E - 1)SD_E^2 + (N_C - 1)SD_C^2}{N_E + N_C - 2}}$$

Tabla 10.13 Elementos fundamentales para interpretar los resultados de una prueba *t*

Estadísticos de grupo										
Género		N	Media		Desviación tip.		Error tip. de la media			
F3	Masculino	86	3.69		1.043		0.113			
	Femenino	88	3.84		1.071		0.114			
Prueba de muestras independientes										
Prueba de Levene para la igualdad de varianzas										
Prueba t para la igualdad de medias										
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% intervalo de confianza para la diferencia	
									inferior	superior
F3	Se han asumido varianzas iguales	0.001	0.970	-0.966	172	0.335	-0.15	0.160	-0.471	0.162
	No se han asumido varianzas iguales			-0.966	171.998	0.335	-0.15	0.160	-0.471	0.162
		Valor "F" diferencia entre las varianzas de los grupos (dispersión de los datos)			Valor "t"		Significancia: no es menor al 0.05, mucho menos al 0.01: No hay diferencias entre los grupos en la variable de contraste			

Donde N_E y N_C son el tamaño de los grupos (grados de libertad), respectivamente; en tanto que, SD_E y SD_C son sus desviaciones estándares.

Ejemplo: $17.9 - 15.2/3.3 = 0.82$ (interpretación: las medias varían menos de una desviación estándar, una respecto de la otra).

Ejemplo: $28.5 - 37.5/4.1 = 2.19$ (los promedios varían más de dos desviaciones estándar uno sobre otro).

¿Qué es la prueba de diferencia de proporciones?

Definición: Es una prueba estadística para analizar si dos proporciones difieren significativamente entre sí.

Hipótesis: De diferencia de proporciones en dos grupos.

Variable: La comparación se realiza sobre una variable. Si hay varias, se efectuará una prueba de diferencia de proporciones por variable.

Nivel de medición de la variable de comparación: Intervalos o razón, expresados en proporciones o porcentajes.

Procedimiento e interpretación: Se obtienen las proporciones de los grupos. Se aplica la siguiente fórmula:

$$z = \frac{P_1 - P_2}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}}$$

$$q_1 = 1 - P_1$$

$$q_2 = 1 - P_2$$

La puntuación z resultante se compara con la puntuación z de la distribución de puntuaciones z (normal) que corresponda al nivel de confianza elegido. El valor calculado de z (resultante de aplicar la fórmula) debe ser igual o mayor que el valor de la tabla de áreas bajo la curva normal correspondiente (tabla 1, apéndice 4, CD anexo). Si es igual o mayor, se acepta la hipótesis de investigación. Si es menor, se rechaza.

EJEMPLO

Hi: El porcentaje de liberales en la ciudad de Aruaim es mayor que en Linderbuck.

% de liberales en Aruaim

% de liberales en Linderbuck

55%

48%

$n_1 = 410$

$n_2 = 301$

Los porcentajes se transforman en proporciones y se calculan q_1 y q_2 :

Aruaim

Linderbuck

$P_1 = 0.55$

$P_2 = 0.48$

$N_1 = 410$

$N_2 = 301$

$q_1 = 1 - .55 = 0.45$

$q_2 = 1 - .48 = 0.52$

$\alpha = .05 = 1.96 z$ (puntuación z que, como se ha explicado anteriormente, corresponde al nivel alfa del .05)

$$z = \frac{0.55 - 0.48}{\sqrt{\frac{(0.55)(0.45)}{410} + \frac{(0.48)(0.52)}{301}}} = 1.849$$

Como la z calculada es menor a 1.96 (nivel alfa expresado en una puntuación z), aceptamos la hipótesis nula y rechazamos la de investigación.

¿Qué es el análisis de varianza unidireccional o de un factor? (*one way*)

Definición: Es una prueba estadística para analizar si más de dos grupos difieren significativamente entre sí en cuanto a sus medias y varianzas. La *prueba t* se utiliza para dos grupos y el *análisis de varianza unidireccional* se usa para tres, cuatro o más grupos. Aunque con dos grupos se puede utilizar también.

Hipótesis: De diferencia entre más de dos grupos. La hipótesis de investigación propone que los grupos difieren significativamente entre sí y la hipótesis nula propone que los grupos no difieren significativamente.

Variables: Una variable independiente y una variable dependiente.

Nivel de medición de las variables: La variable independiente es categórica y la dependiente es por intervalos o razón.

El hecho de que la variable independiente sea categórica significa que es posible formar grupos diferentes. Puede ser una variable nominal, ordinal, por intervalos o de razón (pero en estos últimos dos casos la variable debe reducirse a categorías).

Por ejemplo:

- Religión.
- Nivel socioeconómico (muy alto, alto, medio, bajo y muy bajo).
- Antigüedad en la empresa (de cero a un año, más de un año a cinco años, más de cinco años a 10, más de 10 años a 20 y más de 20 años).

Interpretación: El análisis de varianza unidireccional produce un valor conocido como *F* o razón *F*, que se basa en una distribución muestral, conocida como *distribución F*, que es otro miembro de la familia de distribuciones muestrales. La razón *F* compara las variaciones en las puntuaciones debidas a dos diferentes fuentes: variaciones entre los grupos que se comparan y variaciones dentro de los grupos.

Si los grupos difieren realmente entre sí, sus puntuaciones variarán más de lo que puedan variar las puntuaciones entre los integrantes de un mismo grupo. Veámoslo con un ejemplo cotidiano. Si tenemos tres familias *A*, *B* y *C*. La familia *A* está integrada por Felipe, Angélica, Elena y José Luis. La familia *B* está compuesta por Chester, Pilar, Íñigo, Alonso y Carlos. Y la familia *C* está integrada por Rodrigo, Laura y Roberto. ¿Qué esperamos? Pues que los integrantes de una familia se parezcan más entre sí que a los miembros de otra familia. Esto se graficaría como en la figura 10.23.

Es decir, esperamos *homogeneidad* intrafamilias y *heterogeneidad* interfamilias.

¿Qué sucedería si los miembros de las familias se parecieran más a los integrantes de las otras familias que a los de la suya propia? Quiere decir que no hay diferencia entre los grupos (en el ejemplo, familias).

Esta misma lógica se aplica a la razón *F*, la cual nos indica si las diferencias entre los grupos son mayores que las diferencias intragrupos (dentro de éstos). Estas diferencias se miden en términos de varianza. La *varianza* es una medida de dispersión o variabilidad alrededor de la media y se calcula en términos de desviaciones elevadas al cuadrado. Recuérdese que la *desviación estándar* es un promedio de desviaciones respecto a la media $(X - \bar{X})$ y la *varianza* es un promedio de desviaciones respecto a la media elevadas al cuadrado $(X - \bar{X})^2$. La varianza por eso se simboliza como S^2 y su fórmula es $\Sigma(X - \bar{X})^2/n$. En consecuencia, la razón *F*, que es una razón de varianzas; se expresa así:

Análisis de varianza Prueba estadística para analizar si más de dos grupos difieren entre sí de manera significativa en sus medias y varianzas.

$$F = \frac{\text{Media cuadrática entre los grupos}}{\text{Media cuadrática dentro de los grupos}}$$

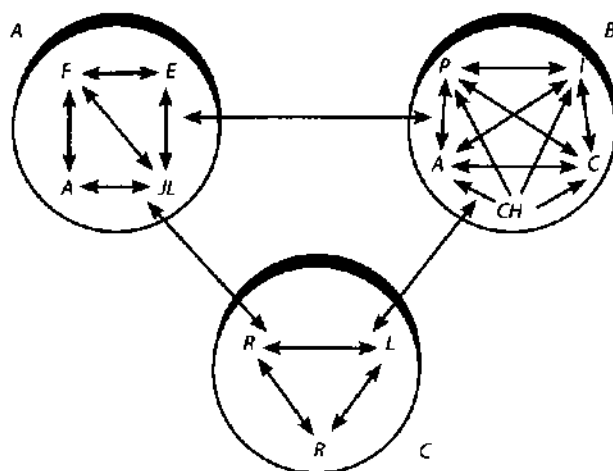


Figura 10.23 Ejemplo de las variaciones de los grupos comparados.

En donde la *media cuadrática* implica un promedio de varianzas elevadas al cuadrado. La media cuadrática entre los grupos se obtiene al calcular la media de las puntuaciones de todos los grupos (media total), después se obtiene la desviación de la media de cada grupo respecto a la media total y se eleva al cuadrado cada una de estas desviaciones, y luego se suman. Finalmente se sopesa el número de individuos en cada grupo y la *media cuadrática* se obtiene con base en los *grados de libertad intergrupales* (no se calcula con base en el número de puntuaciones). La *media cuadrática dentro de los grupos* se calcula obteniendo primero la desviación de cada puntuación respecto a la media de su grupo; posteriormente esta fuente de variación se suma y combina para obtener una media de la *varianza intragrupal* para todas las observaciones, tomando en cuenta los grados de libertad totales (Wright, 1979).

Las fórmulas de la media cuadrática son:

$$\text{Media cuadrática entre grupos} = \frac{\text{Suma de cuadrados entre grupos}}{\text{Grados de libertad entre grupos}}$$

Los grados de libertad entre grupos = $K - 1$ (donde K es el número de grupos).

$$\text{Media cuadrática dentro de los grupos} = \frac{\text{Suma de cuadrados intragrupos}}{\text{Grados de libertad intragrupos}}$$

Los *grados de libertad intragrupos* = $n - K$ (n es el tamaño de la muestra, la suma de los individuos de todos los grupos, y K recordemos que es el número de grupos).

Pues bien, cuando F resulta significativa, quiere decir que los grupos difieren "significativamente" entre sí. Es decir, se acepta la hipótesis de investigación y se rechaza la hipótesis nula.

Cuando se efectúa el *análisis de varianza* por medio de un programa computacional estadístico, se genera una tabla de resultados con los elementos de la tabla 10.14.

El valor *alfa* o probabilidad a elegir una vez más es 0.05 o 0.01. Si es menor que 0.05 es significativo en este nivel, y si es menor que 0.01 también es significativo en este nivel. Cuando el programa o paquete estadístico no incluye la significancia se acude a la tabla 3 del apéndice 4 (tabla de la distribución F). Esta tabla contiene una lista de razones significativas (razones F)

Tabla 10.14 Elementos para interpretar un análisis de varianza unidireccional realizado en SPSS o Minitab

Fuente de variación (Source)	Sumas de cuadrados (Sums of squares)	Grados de libertad (Degrees of freedom)	Medias cuadráticas (Mean squares)	Razón <i>f</i> (<i>f</i> -ratio)	Significancia de <i>f</i> (<i>f</i> prob.)
Entre grupos (between groups)	SS entre	gl entre	SS entre/gl entre	$\frac{M C \text{ entre}}{M C \text{ intra}}$	α
Intragrupos (within groups)	SS intra	gl intra	SS intra/gl intra		
Total	SS entre + SS intra	gl entre + gl intra			

que debemos obtener para aceptar la hipótesis de investigación en los niveles de confianza de 0.05 y 0.01. Al igual que en caso de la razón *t*, el valor exacto de *F* que debemos obtener depende de sus grados de libertad asociados. Por lo tanto, la utilización de la tabla se inicia al buscar los dos valores *gl*, los grados de libertad entre los grupos y los grados de libertad intragrupos. Los grados de libertad entre grupos se indican en la parte superior de la página, mientras que los grados de libertad intragrupos se han colocado al lado izquierdo de la tabla. El cuerpo de la tabla de la distribución *F* presenta razones *F* significativas en los niveles de confianza de 0.05 y 0.01.

$$\text{Si } F = 1.12$$

$$gl \text{ entre} = 2$$

$$gl \text{ intra} = 60$$

Este valor *F* se compara con el valor que aparece en la tabla de la distribución *F* que es 3.15 y como el valor *F* calculado es menor al de dicha tabla, rechazaríamos la hipótesis de investigación y aceptaríamos la hipótesis nula. Para que el valor *F* calculado sea significativo debe ser igual o mayor al de la tabla. A continuación se presenta un ejemplo de un estudio donde el análisis apropiado es el de varianza.

EJEMPLO

- Hi: Los niños que se expongan a contenidos de elevada violencia televisiva exhibirán una conducta más agresiva en sus juegos, respecto de los niños que se expongan a contenidos de mediana o baja violencia televisada.
- Ho: Los niños que se expongan a contenidos de elevada violencia televisiva no exhibirán una conducta más agresiva en sus juegos, respecto de los niños que se expongan a contenidos de mediana o baja violencia televisada.

La variable independiente es el grado de exposición a la violencia televisada y la variable dependiente es la agresividad exhibida en los juegos, medida por el número de conductas agresivas observadas (intervalos).

Para probar la hipótesis se diseña un experimento con cuatro grupos:

$G_1 X_1$ (elevada violencia)	0	} Número de actos agresivos
$G_2 X_2$ (mediana violencia)	0	
$G_3 X_3$ (baja violencia)	0	
G_4 — (conducta prosocial)	0	

En cada grupo hay 25 niños.

La razón F fue de 9.89 y resultó significativa en el nivel de 0.05: se acepta la hipótesis de investigación. La diferencia entre las medias de los grupos es aceptada, el contenido altamente violento tiene un efecto sobre la conducta agresiva de los niños en sus juegos. El estímulo experimental tuvo un efecto. Esto se corrobora comparando las medias de las pospruebas de los cuatro grupos, porque el análisis de varianza unidireccional sólo nos señala si la diferencia entre las medias y las distribuciones de los grupos es o no significativa; pero no nos indica en favor de qué grupos lo es. Es posible hacer esto último al visualizar los promedios y comparar las medias con las distribuciones de los grupos. Y si adicionalmente queremos comparar cada par de medias ($\bar{X}_1$ con $\bar{X}_2$, $\bar{X}_1$ con $\bar{X}_3$, $\bar{X}_2$ con $\bar{X}_3$, etc.) y determinar con exactitud dónde están las diferencias significativas, podemos aplicar un contraste posterior, con el cálculo de una prueba t para cada par de medias; o bien, por medio de algunas estadísticas que suelen ser parte de los análisis efectuados en los paquetes estadísticos computacionales. Tales estadísticas se incluyen en la tabla 10.15.

Tabla 10.15 Principales estadísticas para comparaciones posteriores (*post hoc*) en el ANOVA unidireccional o de un factor

Nombre	Siglas
– Diferencia menos significativa	DMS
– Prueba F de Ryan-Einot-Gabriel-Welsch	R-E-G-W F
– Prueba de rango de Ryan-Einot-Gabriel-Welsch	R-E-G-W Q
– Otras: Waller-Duncan, T2 de Tamhane, T3 de Dunnett, Games-Howell, C de Dunnett, Bonferroni, Sidak, Gabriel, Hochberg...	

Tabla 10.16 Ejemplo del análisis de varianza en SPSS

Fuente de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	Razón f	Significancia de f
Entre grupos	150.18	3	50.06		
Intragrupos	857.64	96	8.93	5.6	0.001
Total	1 007.82	99			

Supongamos que por medio de una escala Likert (1-5) medimos la actitud que tienen hacia el entrenador del equipo de fútbol de una ciudad, las tres porras o grupos de aficionados permanentes: la Ultra, la Central y la de Veteranos. Y queremos analizar si difieren significativamente entre sí. Realizamos el análisis de varianza y los resultados son los que se muestran en la tabla 10.17.

Tabla 10.17 Ejemplo de análisis de varianza

ANOVA

Actitud hacia el entrenador del equipo de fútbol

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	46 768	2	23 384	17 394	0.000
Intra-grupos	793 175	590	1 344		
Total	839 943	592			

Descriptivos

Actitud hacia el entrenador del equipo de fútbol

	N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95%		Mínimo	Máximo
					Límite inferior	Límite superior		
Porra nítida	195	3.61	1.046	0.075	3.46	3.76	1	5
Porra central	208	3.72	1.090	0.076	3.57	3.87	1	5
Porra veteranos	190	3.07	1.331	0.097	3.88	3.26	1	5
Total	593	3.48	1.191	0.049	3.38	3.57	1	5

Comentario: La actitud de las diferentes porras hacia el entrenador es significativamente distinta, la más desfavorable es la de los veteranos (su media es de 3.07, cerrando o redondeando a décimas: 3.1).

Hasta aquí hemos visto pruebas paramétricas con una sola variable independiente y una dependiente. ¿Pero qué ocurre cuando tenemos diversas variables independientes y una dependiente, varias independientes y dependientes? Esquemas del tipo, como se muestra en la figura 10.24.

Entonces, requerimos de otros métodos estadísticos como se muestra en la tabla 10.18.

Estos métodos son comentados en el capítulo 8 del CD adjunto, en análisis multivariado de los datos, junto con el análisis de factores.

... Análisis no paramétricos

¿Cuáles son las presuposiciones de la estadística no paramétrica?

Para realizar los análisis no paramétricos debe partirse de las siguientes consideraciones:

1. La mayoría de estos análisis no requieren de presupuestos acerca de la forma de la distribución poblacional. Aceptan distribuciones no normales.
2. Las variables no necesariamente tienen que estar medidas en un nivel por intervalos o de razón; pueden analizar datos nominales u ordinales. De hecho, si se quieren aplicar análisis no paramétricos a datos por intervalos o razón, éstos se resumen a categorías discretas (a unas cuantas). Las variables deben ser categóricas.

Tabla 10.18 Métodos estadísticos

Método	Propósitos fundamentales
Análisis de varianza factorial (ANOVA de varios factores)	Evaluar el efecto de dos o más variables independientes sobre una variable dependiente.
Análisis de covarianza (ANCOVA)	Analizar la relación entre una variable dependiente y dos o más independientes, al eliminar y controlar el efecto de al menos una de estas variables independientes.
Regresión múltiple	Evaluar el efecto de dos o más variables independientes sobre una variable dependiente, así como predecir el valor de la variable dependiente con una o más variables independientes, y estimar cuál es la independiente que mejor predice las puntuaciones de la variable dependiente. Se trata de una extensión de la regresión lineal.
Análisis multivariado de varianza (MANOVA)	Analizar la relación entre dos o más variables independientes y dos o más variables dependientes.
Análisis lineal de patrones (PATH)	Determinar y representar interrelaciones entre variables a partir de regresiones, así como analizar la magnitud de la influencia de algunas variables sobre otras, influencia directa e indirecta. Es un modelo causal.
Análisis discriminante	Construir un modelo predictivo para pronosticar el grupo de pertenencia de un caso a partir de las características observadas de cada caso (predecir la pertenencia de un caso a una de las categorías de la variable dependiente, sobre la base de dos o más independientes).

¿Cuáles son los métodos o las pruebas estadísticas no paramétricas más utilizados?

Las pruebas no paramétricas más utilizadas son:

1. La *chi* cuadrada o χ^2 .
2. Los coeficientes de correlación e independencia para tabulaciones cruzadas.
3. Los coeficientes de correlación por rangos ordenados de Spearman y Kendall.

Chi cuadrada Prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas.

¿Qué es la *chi* cuadrada o χ^2 ?

Definición: Es una prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas.

Se simboliza: χ^2 .

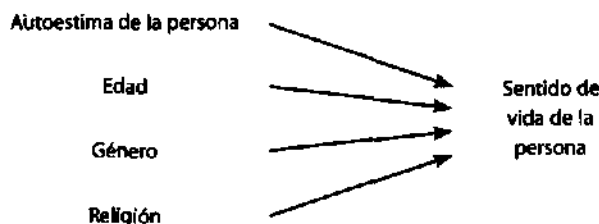
Hipótesis a probar: Correlacionales.

Variables involucradas: Dos. La prueba *chi* cuadrada no considera relaciones causales.

Nivel de medición de las variables: Nominal u ordinal (o intervalos o razón reducidos a ordinales).

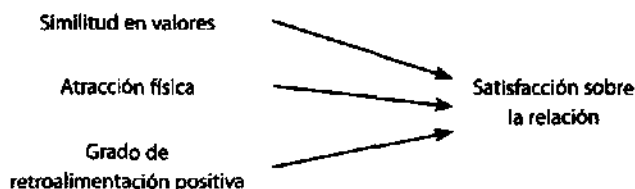
Procedimiento: Se calcula por medio de una *tabla de contingencia* o *tabulación cruzada*, que es un cuadro de dos dimensiones, y cada dimensión contiene una variable. A su vez, cada variable se subdivide en dos o más categorías.

Un ejemplo de una tabla de contingencia se presenta en la tabla 10.19.

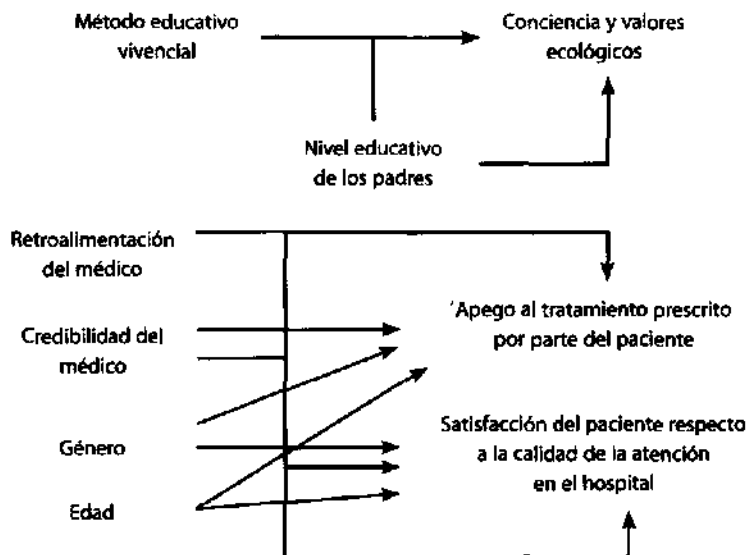


Objetivo: Analizar el efecto que sobre el *sentido de vida* tienen la *autoestima*, la *edad*, el *género* y la *religión*.

O si queremos probar la hipótesis: "La *similitud en valores*, la *atracción física* y el *grado de retroalimentación positiva* son factores que inciden en la *satisfacción sobre la relación* en parejas de novios cuyas edades oscilan entre los 24 y los 32 años."



Asimismo, si pretendemos evaluar si un *método educativo* incrementa la *conciencia y valores ecológicos* de los estudiantes de bachillerato, controlando y analizando la influencia de la variable *nivel educativo de los padres*.



Si buscamos conocer la influencia de cuatro variables de los médicos sobre el apego al tratamiento y la satisfacción en torno a la atención por parte de sus pacientes.

Figura 10.24 Ejemplo de prueba con diversas variables tanto dependientes como independientes.

Tabla 10.19 Ejemplo de una tabla de contingencia

		Voto	
		Candidato A	Candidato B
Sexo	Masculino		
	Femenino		

La tabla 10.19 demuestra el concepto de *tabla de contingencia* o tabulación cruzada. Las variables aparecen señaladas a los lados del cuadro, cada una con sus dos categorías. Se dice que se trata de una tabla 2×2 , donde cada dígito significa una variable y el valor de éste indica el número de categorías de la variable:

$$\begin{array}{ccc}
 2 & \times & 2 \\
 \uparrow & & \uparrow \\
 \text{Una variable con} & & \text{Otra variable} \\
 \text{dos categorías} & & \text{con dos categorías}
 \end{array}$$

Un ejemplo de una tabla de contingencia 2×3 se muestra en la tabla 10.20.

En la tabla de contingencia se anotan las *frecuencias observadas* en la muestra de la investigación, como ocurre en la tabla 10.20.

Tabla 10.20 Ejemplo de una tabla de contingencia 2×3

		Zona del distrito electoral		Total
		Norte	Sur	
Identificación política	Partido Derechista	180	100	280
	Partido del Centro	190	280	470
	Partido Izquierdista	170	120	290
Total		540	500	1040

Después, se calculan las *frecuencias esperadas* para cada celda. En esencia, la *chi cuadrada* es una *comparación* entre la *tabla de frecuencias observadas* y la denominada *tabla de frecuencias esperadas*, la cual constituye la tabla que esperaríamos encontrar si las variables fueran estadística-

mente independientes o no estuvieran relacionadas (Wright, 1979). La *chi cuadrada* es una prueba que parte del supuesto de "no relación entre variables" y el investigador evalúa si en su caso esto es cierto o no, analiza si las frecuencias observadas son diferentes de lo que pudiera esperarse en caso de ausencia de correlación. La lógica es así: "Si no hay relación entre las variables, debe tenerse una tabla así (el de las frecuencias esperadas). Si hay relación, la tabla que obtengamos como resultado en nuestra investigación tiene que ser muy diferente respecto de la tabla de frecuencias esperadas."

La frecuencia esperada de cada celda, casilla o recuadro, se calcula mediante la siguiente fórmula aplicada a la tabla de frecuencias observadas.

$$fe = \frac{(\text{Total o marginal de renglón})(\text{total o marginal de columna})}{n}$$

En donde n es el número total de frecuencias observadas.

Para la primera celda (zona norte y partido derechista) la frecuencia esperada sería:

$$fe = \frac{(280)(540)}{1040} = 145.4$$

Veamos de dónde salieron los números:

		280 total de renglón
540 Total de columnas		1040 n

Para el ejemplo de la tabla 10.20, la tabla de frecuencias esperadas sería la tabla 10.21.

Tabla 10.21 Cuadro de frecuencias esperadas para la tabla 10.20

145.4	134.6	280
244.0	226.0	470
150.6	139.4	290
540	500	1040

Una vez obtenidas las frecuencias esperadas, se aplica la siguiente *fórmula de chi cuadrada*:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

donde Σ significa sumatoria.

O es la frecuencia observada en cada celda.

E es la frecuencia esperada en cada celda.

Es decir, se calcula para cada celda la diferencia entre la frecuencia observada y la esperada; esta diferencia se eleva al cuadrado y se divide entre la frecuencia esperada. Finalmente, se suman tales resultados y la sumatoria es el valor de χ^2 obtenida.

Otra manera de calcular χ^2 es mediante la tabla 10.22.

Tabla 10.22 Procedimiento para calcular la *chi* cuadrada

Celda	O	E	$O - E$	$(O - E)^2$	$\frac{(O - E)^2}{E}$
Zona norte/partido derechista	180	145.4	34.6	1 197.16	8.23
Zona norte/partido centro	190	244.4	-54.4	2 959.36	12.11
Zona norte/partido izquierdista	170	150.6	19.4	376.36	2.50
Zona sur/partido derechista	100	134.6	-34.6	1 197.16	8.89
Zona sur/partido centro	280	226.0	54.0	2 916.00	12.90
Zona sur/partido izquierdista	120	139.4	-19.4	376.36	2.70
					$\chi^2 = 47.33$

El valor χ^2 para los valores observados es de 47.33.

Interpretación: Al igual que t y F , la *chi cuadrada* proviene de una distribución muestral denominada distribución χ^2 , y los resultados obtenidos en la muestra están identificados por los grados de libertad. Esto es, para saber si un valor de χ^2 es o no significativo, debemos calcular los grados de libertad. Éstos se obtienen mediante la siguiente fórmula:

$$gl = (r - 1)(c - 1)$$

En donde r es el número de renglones del cuadro de contingencia y c el número de columnas. En nuestro caso:

$$gl = (3 - 1)(2 - 1) = 2$$

Y acudimos con los grados de libertad que corresponden a la tabla 4 del apéndice 4 del CD (Distribución de χ^2 cuadrada), eligiendo nuestro nivel de confianza (0.05 o 0.01). Si nuestro valor calculado de χ^2 es igual o superior al de la tabla, decimos que las variables están relacionadas (χ^2 fue significativa). En el ejemplo, el valor que requerimos empatar o superar al nivel del 0.05 es 5.991. El valor de χ^2 calculado por nosotros es de 47.33, que es muy superior al de la tabla: χ^2 resulta significativa.

EJEMPLO

Hi: Los tres canales de televisión a nivel nacional difieren en la cantidad de programas prosociales, neutrales y antisociales que difunden. Hay relación entre la variable 'canal de televisión nacional' y la variable 'emisión de programas prosociales, neutrales y antisociales'.

Resultados: $\chi^2 = 7.95$
 $gl = 4$

Para que χ^2 sea significativa a 0.01, con cuatro grados de libertad, se necesita un valor mínimo de 13.277 y para que sea significativa a 0.05, se necesita un valor mínimo de 9.488. Por lo tanto, se rechaza la hipótesis de investigación y se acepta la nula. No hay relación entre las variables.

Comentario: Cuando al calcular χ^2 se utiliza un paquete estadístico computacional, el resultado de *chi cuadrada* se proporciona junto con su significancia; si ésta es menor a 0.05 o a 0.01, se acepta la hipótesis de investigación.

Tabla 10.23 Tabla de contingencia correspondiente al ejemplo

Esquema:	Canal 24	Canal 23	Canal 56
Programas prosociales			
Programas neutrales			
Programas antisociales			

¿Qué son los coeficientes de correlación e independencia para tabulaciones cruzadas?

Adicionales a la *chi cuadrada*, existen otros coeficientes para evaluar si las variables incluidas en la tabla de contingencia o tabulación cruzada están correlacionadas. En la tabla 10.24 se describen los coeficientes más importantes para tal finalidad.

Tabla 10.24 Principales coeficientes para tablas de contingencia

Coeficiente	Para cuadros de contingencia...	Nivel de medición de ambas variables	Interpretación
<i>Phi</i> (ϕ)	2×2	Nominal. Puede utilizarse con variables ordinales reducidas a dos categorías. SPSS lo muestra en cálculos para datos nominales.	En tablas 2×2 varía de 0 a 1, donde cero implica ausencia de correlación entre las variables; y uno, que hay correlación perfecta entre las variables. En tablas más grandes, <i>phi</i> puede ser mayor de 1.0, pero la interpretación es compleja. Por ello, se recomienda limitar su uso a las tablas 2×2 .
Coeficiente de contingencia <i>C</i> de Pearson	Cualquier tamaño. De hecho es un ajuste de <i>phi</i> para tablas con más de dos categorías en las variables. Incluso funciona mejor con tablas de 5×5 .	Nominal. Puede utilizarse con variables ordinales reducidas a dos categorías.	0 a 1, pero en tablas menores a 5×5 , se acerca pero nunca alcanza el uno.
V de Cramer (<i>C</i>)	Cualquier tamaño	Cualquier nivel de variables, pero siempre reducidas a categorías. SPSS lo muestra en cálculos para datos nominales.	0 a 1, pero el uno solamente se alcanza si ambas variables tienen el mismo número de categorías (o marginales).
Goodman-Kruskal <i>Lambda</i> o sólo <i>Lambda</i> (λ)	Cualquier tamaño	Cualquier nivel de variables, pero siempre reducidas a categorías. SPSS lo muestra en cálculos para datos nominales.	Fluctúa entre 0 y 1, asume causalidad, lo que significa que puede predecirse a la variable dependiente definida en la tabla, sobre la base de la independiente. La versión usual de <i>Lambda</i> es asimétrica. Sin embargo, SPSS y otros programas presentan tres versiones: una simétrica y dos asimétricas (estas últimas representan a cada una de las variables considerada como dependiente). La versión simétrica es simplemente el promedio de las dos <i>Lambdas</i> .

(continúa)

Tabla 10.24 Principales coeficientes para tablas de contingencia (*continuación*)

Coeficiente	Para cuadros de contingencia...	Nivel de medición de ambas variables	Interpretación
			asimétricas. Una prueba asimétrica presupone que el investigador puede designar cuál es la variable independiente y cuál la dependiente. En una simétrica no se asume tal causalidad.
Coeficiente de incertidumbre o entropía o <i>U</i> de Theil	Cualquier tamaño	Cualquier nivel de variables, pero siempre reducidas a categorías. SPSS lo muestra en en cálculos para datos nominales.	Fluctúa entre 0 y 1, asume causalidad, lo que significa que puede predecirse a la variable dependiente definida en la tabla, sobre la base de la independiente. Por razones históricas (de costumbre), el coeficiente se ha computado frecuentemente en términos de predecir la variable de las columnas, sobre la base de la variable de las filas.
<i>Gamma</i> de Goodman y Kruskal	Cualquier tamaño	Ordinal	Varía de -1 a $+1$ (-1 es una relación negativa perfecta, y $+1$ una relación positiva perfecta).
<i>Tau-a</i> , <i>Tau-b</i> y <i>Tau-c</i> (τ_a , τ_b , τ_c)	Cualquier tamaño	Ordinal	Varían de -1 a $+1$. <i>Tau-a</i> y <i>Tau-b</i> son asimétricas y <i>Tau-c</i> es simétrica.
<i>D</i> de Somers	Cualquier tamaño	Ordinal	Varía de -1 a $+1$.
<i>Kappa</i>	Cualquier tamaño	Datos categorizados por intervalo.	Regularmente de 0 a 1.

¿Qué otra aplicación tienen las tablas de contingencia?

Las tablas de contingencia, además de servir para el cálculo de *chi cuadrada* y otros coeficientes, son útiles para describir conjuntamente dos o más variables. Esto se efectúa al convertir las frecuencias observadas en frecuencias relativas o porcentajes. En una tabulación cruzada puede haber tres tipos de porcentajes respecto de cada celda.

- Porcentaje en relación con el total de frecuencias observadas ("N" o "n" de muestra).
- Porcentaje en relación con el total marginal de la columna.
- Porcentaje en relación con el total marginal del renglón.

Veamos con un ejemplo hipotético de una tabla 2×2 con las variables género y preferencia por un conductor televisivo. Las frecuencias observadas serían:

EJEMPLO

		Género		
		masculino	femenino	
Preferencia por el conductor	A	25	25	50
	B	40	10	50
		65	35	100

Las celdas podrían representarse así:

a	c
b	d

Tomemos el caso de a (celda superior izquierda). La celda a (25 frecuencias observadas) con respecto al total ($N = 100$) representa 25%. En relación con el total marginal de columna (cuyo total es 65) representa 38.46% y respecto del total marginal de renglón (cuyo total es 50) significa 50%. Esto puede expresarse así:

EJEMPLO

		Frecuencias observadas	
En relación con N	25		$a + c = 50$
En relación con " $a + b$ "	25.00%	c	
En relación con " $a + c$ "	38.46%		
	50.00%		
	b	d	$b + d$
	$a + b = 65$	$c + d$	$100 = N$

Así procedemos con cada categoría, como ocurre en la tabla 10.25.

Comentario: Una cuarta parte de la muestra está constituida por hombres que prefieren al conductor A, 10.0% son mujeres que prefieren al conductor B. Más de 60% (61.5%) de los hombres al B, etcétera.

Tabla 10.25 Ejemplo de una tabla de contingencia para describir conjuntamente dos variables

Preferencia por el conductor		Sexo		
		Masculino	Femenino	
A	25	25	50	
	25.0%	25.0%		
	38.5%	71.4%		
	50.0%	50.0%		
B	40	10	50	
	40.0%	10.0%		
	61.5%	28.6%		
	80.0%	20.0%		
		65	35	100

Debe observarse que tales *porcentajes* se basan en las *frecuencias observadas*, pero no tienen nada que ver con las *frecuencias esperadas* (estas últimas son frecuencias absolutas). Además, para evaluar la fuerza de la relación entre las dos variables disponemos de los coeficientes para tablas de contingencia mencionados, cuya significancia se interpreta como *chi cuadrada* y como cualquier resultado de los revisados en este capítulo.

Otros coeficientes de correlación

El **coeficiente de correlación de Pearson** es una estadística apropiada para variables medidas por intervalos o razón y para relaciones lineales. La *chi cuadrada* y demás coeficientes son estadísticas adecuadas para tablas de contingencia con variables nominales, ordinales y de intervalos, pero reducidas a categorías ¿pero qué ocurre si las variables de nuestro estudio son ordinales, por intervalos y de razón?, o bien, una mezcla de niveles de medición, o los datos no necesariamente los disponemos en una tabla de contingencia. Existen otros coeficientes que comentaremos brevemente.

Coeficiente de correlación de Pearson Es una estadística apropiada para variables medidas por intervalos o razón y para relaciones lineales.

¿Qué son los coeficientes y la correlación por rangos ordenados de Spearman y Kendall?

Los **coeficientes rho de Spearman**, simbolizado como r_s , y **tau de Kendall**, simbolizado como t , son medidas de correlación para variables en un nivel de medición ordinal (ambas), de tal modo que los individuos u objetos de la muestra pueden ordenarse por rangos (jerarquías). Por ejemplo, supongamos que tenemos las variables "preferencia en el sabor" y "atractivo del envase", y

Coeficientes rho de Spearman y tau de Kendall Son medidas de correlación para variables en un nivel de medición ordinal; los individuos u objetos de la muestra pueden ordenarse por rangos.

pedimos a un grupo de personas representativas del mercado que evalúen conjuntamente 10 refrescos embotellados específicos y los ordenen del 1 al 10; en tanto que, "1" es la categoría o el rango máximo en ambas variables. Finalmente se obtienen los siguientes resultados del grupo:

<i>Refresco⁷</i>	<i>Variable 1 Preferencia en el sabor</i>	<i>Variable 2 Atractivo del envase</i>
Loy	1	2
Caravana	2	5
Mauna-Loa	3	1
Recreo	4	3
Carma	5	4
Manzanol	6	6
Cereza	7	8
Pezcara	8	7
Casa	9	10
Manzanita	10	9

Para analizar tales resultados, utilizaríamos los coeficientes r_s y t . Ahora bien, debe observarse que todos los refrescos o sodas tienen que jerarquizarse por rangos que contienen las propiedades de una escala ordinal (se ordena de mayor a menor). Ambos coeficientes varían de -1.0 (correlación negativa perfecta) a $+1.0$ (correlación positiva perfecta), considerando el 0 como ausencia de correlación entre las variables jerarquizadas. Se trata de estadísticas sumamente eficientes para datos ordinales. La diferencia entre ellos es explicada por Nie *et al.* (1975, p. 289) de la manera siguiente: El coeficiente de Kendall (t) resulta un poco más significativo cuando los datos contienen un número considerable de rangos empatados. El coeficiente de Spearman ρ parece ser una aproximación cercana al coeficiente r de Pearson, cuando los datos son continuos (por ejemplo, no caracterizados por un número considerable de empates en cada rango). De acuerdo con Creswell (2005) sirve también para analizar relaciones curvilíneas.

También se interpreta su significancia. Otro ejemplo sería relacionar la opinión de dos médicos y la jerarquización de los mismos pacientes en cuanto al avance de una enfermedad terminal en éstos.

¿Qué otros coeficientes existen?

Un coeficiente muy importante es el η^2 , que es similar al coeficiente r de Pearson, pero con relaciones no lineales, las cuales se comentaron anteriormente. Es decir, η^2 define la "correlación perfecta" (1.00) como curvilínea y a la "relación nula" (0.0) como la independencia estadística de las variables. Este coeficiente es asimétrico (concepto explicado en la tabla 10.25), y a diferencia de Pearson, se puede obtener un valor diferente para el coeficiente; al determinar cuál variable se considera independiente y cuál dependiente. η^2 es interpretada como el porcentaje de la varianza en la variable dependiente explicado por la independiente. El investigador puede calcular η^2 de las dos maneras: al cambiar la definición de la independiente y dependiente, luego

⁷ Nombres ficticios.

promediar los dos coeficientes y obtener uno simétrico. *Eta* puede trabajarse en tablas de contingencia. Otros coeficientes se describen en la tabla 10.26.

Tabla 10.26 Otros coeficientes

Coeficiente	Nivel de medición de las variables	Ejemplo	Interpretación
Biserial (r_y)	Una ordinal y la otra por intervalos o razón.	Jerarquía en la organización y motivación.	-1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).
Biserial por rangos (r_{rb})	Una variable ordinal y la otra nominal.	Escuela de procedencia (pública-privada) y rango en una prueba de un idioma extranjero (alto, medio, bajo).	-1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).
Biserial puntual (r_{pb})	Una variable nominal y la otra por intervalos o razón.	Motivación al estudio y licenciatura (Economía, Derecho, Administración, etcétera).	-1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).
<i>Tetrachoric</i>	Las dos dicotómicas, no necesariamente expresadas en tablas.	Género y afiliación/no afiliación a un partido político.	-1.00 (correlación negativa perfecta). 0.0 (ausencia de relación) +1.00 (correlación positiva perfecta).

Existen muchos más coeficientes, pero los más importantes son los señalados. Lo mejor de todo es que los programas computacionales de análisis estadístico los calculan, lo único que tenemos que hacer es interpretarlos y verbalizar sus resultados con comentarios. Veamos algunos ejemplos de secuencias.

En SPSS

Prueba *t*:

Analizar → Comparar medias → Prueba T

Coeficientes para tablas de contingencia:

Analizar → Estadísticos descriptivos → Estadísticos → *Phi*, *Lambda*, etc.

En Minitab

Análisis de varianza unidireccional:

Stat → ANOVA → *One-Way*

Aunque en el futuro cambien los nombres en los comandos, el proceso será similar, por lo que rápidamente nos familiarizaremos con ellos.

Ahora, presentamos un cuadro final sobre los principales métodos estadísticos, considerando: a) *el tipo de pregunta de investigación* (descriptiva, de diferencia de grupos, correlacional o causal), b) *el número de variables involucradas*, c) *nivel de medición de las variables o tipo de datos* y d) *en comparación de grupos*, si son *muestras independientes* o *correlacionadas*. En este último punto, las muestras independientes se seleccionan de manera que no exista ninguna relación entre los miembros de las muestras; por ejemplo, un grupo experimental y uno de control en un experimento. No hay ningún emparejamiento de las observaciones entre las muestras. Mientras que en las correlacionadas, sí existe una relación entre los miembros de las muestras; por ejemplo, el mismo grupo antes y después de un tratamiento experimental, preprueba y posprueba. Algunas de las pruebas o métodos estadísticos no fueron desarrollados en el capítulo y varios se encuentran en el capítulo 8 del CD anexo.

Tabla 10.27⁸ Elección de los procedimientos estadísticos o pruebas

1) Pregunta de investigación: Descriptiva	Procedimiento o prueba
Datos nominales	Moda
Datos ordinales	Mediana, moda
Datos por intervalos o razón	Media, mediana, moda, desviación estándar, varianza y rango
2) Pregunta de investigación: Diferencias de grupos	
a) Dos variables (grupos)	
a. 1. Muestras correlacionadas	
• Datos nominales	Prueba de McNemar
• Datos ordinales	Prueba de Wilcoxon para pares de rangos
• Datos por intervalos o razón	Prueba <i>t</i> para muestras correlacionadas
a. 2. Muestras independientes	
• Datos nominales	<i>Chi</i> cuadrada
• Datos ordinales	Prueba Mann-Whitney U o prueba Kolmogorov-Smirnov para dos muestras
• Datos por intervalos o razón	Prueba <i>t</i> para muestras no correlacionadas o independientes
b) Más de dos variables (grupos)	
b. 1. Muestras correlacionadas	
• Datos nominales	Prueba Q de Cochran
• Datos ordinales	Análisis de varianza de Friedman en dos vías
• Datos por intervalos o razón	Análisis de varianza (ANOVA)
• Datos por intervalos o razón, control de efectos de otra variable independiente	Análisis de covarianza (ANCOVA)

(continúa)

⁸ Adaptado de Mertens (2005, p. 409).

Tabla 10.27 Elección de los procedimientos estadísticos o pruebas (*continuación*)

b. 2. Muestras independientes	
• Datos nominales	Chi cuadrada para k muestras independientes
• Datos nominales u ordinales (categóricos) y de intervalos-razón	Chi cuadrada de Friedman
• Datos ordinales	Análisis de varianza en una vía de Kruskal-Wallis (ANOVA)
• Datos por intervalos o razón	Análisis de varianza (ANOVA)
3) Pregunta de investigación:	
Correlacional	
a) Dos variables	
• Datos nominales	Coefficiente de contingencia o Φ
• Datos ordinales	Coefficiente de rangos ordenados de Spearman o coeficiente de rangos ordenados de Kendall
• Datos por intervalos o razón	Coefficiente de correlación de Pearson (producto-momento)
• Una variable independiente y una dependiente (ambas de intervalos o razón)	Regresión lineal
• Datos por intervalos y nominales u ordinales	Coefficiente biserial puntual
• Datos por intervalos y una dicotomía artificial en una escala ordinal (la dicotomía es artificial porque subyace una distribución continua)	Coefficiente biserial
b) Más de dos variables	
• Datos nominales	Análisis discriminante
• Datos ordinales	Análisis de correlación parcial por rangos de Kendall
• Datos por intervalos o razón	Coefficiente de correlación parcial o múltiple, R^2
4) Pregunta de investigación:	
Causal o predictiva	
• Diversas independientes y una dependiente (las independientes en cualquier nivel de medición, la dependiente en nivel por intervalos o razón). Cuando las independientes son nominales u ordinales se convierten en variables "dummy"	Regresión múltiple
Diversas independientes y dependientes	Análisis multivariado de varianza (MANOVA)
• Agrupamiento (membresía de todos los datos)	Análisis discriminante (en una vía, jerárquico o factorial, de acuerdo con el número de variables involucradas)
• Estructuras y redes causales	Análisis de patrones o vías (<i>path analysis</i>)
5) Pregunta de investigación:	
Estructura de variables o validación de constructo	
Las variables deben estar por intervalos o razón	
	Análisis de factores

Paso 6: realizar análisis adicionales

Este paso implica —simplemente— que una vez realizados nuestros análisis, es posible que decidamos realizar otros análisis extras o pruebas para confirmar tendencias y evaluar los datos desde diferentes ángulos. Por ejemplo, podemos en una tabla de contingencia calcular primero *Chi* cuadrada y luego *Phi*, *Lambda*, *T* de Cramer (C) y el coeficiente de contingencia. O después de un ANOVA, efectuar los contrastes posteriores que consideremos apropiados. Resulta este paso un momento clave para verificar que no se nos haya olvidado un análisis pertinente.

Paso 7: preparar los resultados para presentarlos

Se recomienda, una vez que se obtengan los resultados de los análisis estadísticos (tablas, gráficas, cuadros, etc.), las siguientes actividades —sobre todo para quienes se inician en la investigación—:

1. Revisar cada resultado [análisis general → análisis específico → valores resultantes (incluida la significancia) → tablas, diagramas, cuadros y gráficas].
2. Organizar los resultados (primero los descriptivos, por variable; luego los resultados relativos a la confiabilidad y la validez; posteriormente los inferenciales, que se pueden ordenar por hipótesis o de acuerdo con su desarrollo).
3. Cotejar diferentes resultados: su congruencia y en caso de inconsistencia lógica volverlos a revisar. Asimismo, se debe evitar la combinación de tablas, diagramas o gráficas que repitan datos. Por lo común, columnas o filas idénticas de datos no deben aparecer en dos o más tablas. Cuando éste es el caso, debemos elegir la tabla o elemento que ilustre o refleje mejor los resultados y sea la opción que presente mayor claridad. Una buena pregunta en este momento del proceso es: ¿qué valores, tablas, diagramas, cuadros o gráficas son necesarios?, ¿cuáles explican mejor los resultados?
4. Priorizar la información más valiosa (que es en gran parte resultado de la actividad anterior), sobre todo si se van a producir reportes ejecutivos y otros más extensos.
5. Copiar y/o “formatear” las tablas en el programa con el cual se elaborará el reporte de la investigación (procesador de textos o uno para presentaciones, como Word o Power Point). Algunos programas como SPSS y Minitab permiten que se transfieran los resultados (tablas, por ejemplo) directamente a otro programa (copiar y pegar). Por ello, resulta conveniente usar una versión del programa de análisis que esté en el mismo idioma del programa que se empleará para escribir el reporte o elaborar la presentación.
6. Comentar o describir brevemente la esencia de los análisis, valores, tablas, diagramas, gráficas.
7. Volver a revisar los resultados.
8. Y, finalmente, elaborar el reporte de investigación.



- El análisis cuantitativo de los datos se efectúa mediante la matriz de datos, la cual está guardada en un archivo.
- Los pasos más importantes en el análisis de los datos son:
 - Decidir el programa de análisis de los datos a utilizar.
 - Explorar los datos obtenidos en la recolección:
 - a) Analizar descriptivamente los datos por variable.
 - b) Visualizar los datos por variable.
 - Evaluar la confiabilidad y validez del instrumento de medición utilizado.
 - Analizar e interpretar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial).
 - Realizar análisis adicionales.
 - Preparar los resultados para presentarlos.
- Los análisis estadísticos se llevan a cabo mediante programas computacionales, con la ayuda de paquetes estadísticos, los más conocidos son: SPSS, Minitab y SAS.
- El tipo de análisis o pruebas estadísticas depende del nivel de medición de las variables, las hipótesis y el interés del investigador.
- Los principales análisis estadísticos que pueden hacerse son: estadística descriptiva para cada variable (distribución de frecuencias, medidas de tendencia central y medidas de la variabilidad), la transformación a puntuaciones z , razones y tasas, cálculos de estadística inferencial, pruebas paramétricas, pruebas no paramétricas y análisis multivariados.
- Las distribuciones de frecuencias contienen las categorías, los códigos, las frecuencias absolutas (número de casos), los porcentajes, los porcentajes válidos y los porcentajes acumulados.
- Las distribuciones de frecuencias (particularmente de los porcentajes) pueden presentarse en forma gráfica.
- Una distribución de frecuencias puede representarse por medio del polígono de frecuencias o de la curva de frecuencias.
- Las medidas de tendencia central son la moda, la mediana y la media.
- Las medidas de la variabilidad son el rango (diferencia entre el máximo y el mínimo), la desviación estándar y la varianza.
- Otras estadísticas descriptivas de utilidad son la asimetría y la curtosis.
- Las puntuaciones z son transformaciones de los valores obtenidos a unidades de desviación estándar.
- Una razón es la relación entre dos categorías; una tasa es la relación entre el número de casos de una categoría y el número total de casos, multiplicada por un múltiplo de diez.
- La confiabilidad se calcula mediante coeficientes: de correlación, *alfa* y KR-20 y 21.
- La validez de criterio se obtiene mediante coeficientes de correlación y la de constructo por medio del análisis de factores.
- La estadística inferencial sirve para efectuar generalizaciones de la muestra a la población. Se utiliza para probar hipótesis y estimar parámetros. Se basa en el concepto de distribución muestral.
- La curva o distribución normal es un modelo teórico sumamente útil; su media es 0 (cero) y su desviación estándar es uno (1).

- El nivel de significancia y el intervalo de confianza son niveles de probabilidad de cometer un error, o de equivocarse en la prueba de hipótesis o la estimación de parámetros. Los niveles más comunes son 0.05 y 0.01.
- Los análisis o las pruebas estadísticas paramétricas más utilizados son:

Prueba	Tipo de hipótesis
Coeficiente de correlación de Pearson	Correlacional
Regresión lineal	Correlacional/causal
Prueba <i>t</i>	Diferencia de grupos
Contraste de la diferencia de proporciones	Diferencia de grupos
Análisis de varianza (ANOVA): unidireccional con una variable independiente y factorial con dos o más variables independientes	Diferencia de grupos/causal
Análisis de covarianza (ANCOVA)	Correlacional/causal

- En todas las pruebas estadísticas paramétricas las variables están medidas en un nivel por intervalos o razón.
- Los análisis o las pruebas estadísticas no paramétricas más utilizados son:

Prueba	Tipos de hipótesis
<i>Chi</i> cuadrada	Diferencias de grupos para establecer correlación
Coeficiente de correlación e independencia para tabulaciones cruzadas: phi, <i>C</i> de Pearson, <i>V</i> de Cramer, lambda, gamma, tau (varios), Somers, etcétera	Correlacional
Coeficientes de correlación de Spearman y Kendall	Correlacional
Coeficiente eta para relaciones no lineales (ejemplos: curvilineales)	Correlacional

- Las pruebas no paramétricas se utilizan con variables nominales u ordinales.
- Los análisis multivariados trabajan con más de una variable de manera simultánea y se presentan en el capítulo 8 del CD anexo.
- Una vez analizados los datos, los resultados se preparan para incluirse en el reporte de la investigación.



CONCEPTOS BÁSICOS

Análisis de datos
 Análisis de factores
 Análisis de varianza
 Análisis multivariados
 Asimetría
 Categoría
 Chi cuadrada
 Codificación
 Coeficiente de correlación de Pearson
 Coeficiente de Kendall
 Coeficiente de Spearman
 Coeficientes de correlación e independencia para tabulaciones cruzadas
 Contraste de diferencia de proporciones
 Curtosis
 Curva de frecuencias
 Curva o distribución normal
 Desviación estándar
 Distribución de frecuencias
 Estadística
 Estadística descriptiva
 Estadística inferencial
 Estadística no paramétrica
 Estadística paramétrica
 Eta
 Gráficas
 Intervalo de confianza
 Matriz de datos
 Media
 Mediana
 Medida de tendencia central
 Medidas de variabilidad
 Métodos cuantitativos
 Minitab
 Moda
 Nivel de significancia
 Paquetes estadísticos
 Polígono de frecuencias
 Prueba *t*
 Pruebas estadísticas
 Puntuación *z*
 Rango
 Razón
 Regresión lineal
 Tabulación cruzada
 Tasa
 Variable de la matriz de datos
 Variable del estudio
 Varianza

EJERCICIOS

1. Construya una distribución de frecuencias hipotéticas, con todos sus elementos, e interprétela verbalmente.
2. Localice una investigación científica donde se reporte la estadística descriptiva de las variables y analice las propiedades de cada estadígrafo o información estadística proporcionada (distribución de frecuencias, medidas de tendencia central y medidas de la variabilidad).
3. Un investigador obtuvo, en una muestra, las siguientes frecuencias abso-

lutas para la variable "actitud hacia el director de la escuela":

<i>Categoría</i>	<i>Frecuencias absolutas</i>
Totalmente desfavorable	69
Desfavorable	28
Ni favorable ni desfavorable	20
Favorable	13
Totalmente favorable	6

- a) Calcule las frecuencias relativas o porcentajes.

- b) Grafique los porcentajes mediante un histograma (barras).
- c) Explique los resultados para responder a la pregunta: ¿la actitud hacia el director de la escuela tiende a ser favorable o desfavorable?
4. Un investigador obtuvo, en una muestra de trabajadores, los siguientes resultados al medir el "orgullo por el trabajo realizado". La escala oscilaba entre 0 (nada de orgullo por el trabajo realizado) a 8 (orgullo total).
- Máximo = 5
 Mínimo = 0
 Media = 3.6
 Moda = 3.0

Mediana = 3.2

Desviación estándar = 0.6

¿Qué puede decirse en esta muestra acerca del orgullo por el trabajo realizado?

5. ¿Qué es la curva normal? ¿Qué son el nivel de significancia y el intervalo de confianza? Responda a estas preguntas en equipo con sus compañeros.
6. Relacione las columnas A y B. En la columna A se presentan hipótesis; y en la columna B, pruebas estadísticas apropiadas para las hipótesis. Se trata de encontrar la prueba que corresponde a cada hipótesis (las respuestas se localizan en el CD anexo: "Respuestas a los ejercicios").

Columna A

- Hi: A mayor inteligencia, mayor capacidad de resolver problemas matemáticos (medidas las variables por intervalos).
- Hi: Los hijos de padres alcohólicos muestran una menor autoestima con respecto a los hijos de padres no alcohólicos (autoestima medida por intervalos).
- Hi: El porcentaje de delitos por asalto a mano armada, en relación con el total de crímenes cometidos, es mayor en la ciudad de México que en Caracas.
- Hi: El sexo está relacionado con la preferencia por telenovelas o espectáculos deportivos.

Columna B

- Diferencias de proporciones
- *Chi* cuadrada
- Spearman
- Coeficiente de correlación de Pearson

- Hi: La intensidad del sabor de productos empacados de pescado está relacionada con la preferencia por la marca (sabor intenso, sabor medianamente intenso, sabor poco intenso, sabor muy poco intenso) (preferencia = rangos a 12 marcas). — ANOVA unidireccional
- Hi: Se presentarán diferencias en cuanto al aprovechamiento entre un grupo expuesto a un método de enseñanza novedoso, un grupo que recibe instrucción mediante un método tradicional y un grupo de control que no se expone a ningún método. — Prueba *t*

7. Un investigador obtuvo un valor *t* igual a 3.25, con 63 grados de libertad y un nivel de confianza o significancia de 0.05, ¿aceptará su hipótesis de investigación? (Respuesta en el CD anexo: Respuestas a los ejercicios.)
8. Otro investigador obtuvo un valor de χ^2 (*chi* cuadrada) de 6.12, con tres grados de libertad y un nivel *alfa* de 0.05, ¿aceptará su hipótesis de investigación? (Respuesta en el CD anexo: Respuestas a los ejercicios.)
9. Suponga un estudio cuya variable independiente es: años de experiencia del docente, y la dependiente: satisfacción del grupo (ambas medidas por intervalos), ¿qué pruebas y modelo estadístico le servirían para analizar los datos y cómo podrá efectuarse el análisis?
10. Genere un ejemplo hipotético de una razón "*F*" significativa e interprete-la.
11. Construya un ejemplo hipotético de una tabulación cruzada y utilícela para fines descriptivos.
12. Busque un artículo de investigación social en revistas científicas que contengan resultados de pruebas *t*, ANOVA y χ^2 aplicadas; evalúe la interpretación de los autores.
13. Para interpretar una prueba se requiere evaluar el resultado (valor) y..... (complete la frase).
14. Respecto al estudio que ha ido desarrollando a lo largo del proceso cuantitativo, ¿qué pruebas estadísticas le serán útiles para analizar los datos? y ¿qué secuencia de análisis habrá de seguir? (Discútalo con su profesor y sus compañeros).

Comentario: Se presentan análisis sencillos y algunos complejos, algunos visuales y otros más al estilo de las revistas científicas. Cada lector visualizará las partes que le resulten útiles. Con esta variedad pretendemos ampliar el marco de los ejemplos. Desde luego, por cuestiones de espacio no se incluyen todos los resultados. Se agrega un ejemplo más cotidiano.

LA TELEVISIÓN Y EL NIÑO

El análisis se realizó utilizando el Minitab. Los datos son diversos para incluirlos en este espacio, incluimos únicamente la gráfica (ver figura 10.26) de sus programas preferidos en 2005.

Comentario: Los tres favoritos: Bob Esponja, las telenovelas infantiles y Los Simpson(s).

EL CLIMA ORGANIZACIONAL

Los análisis del estudio se llevaron a cabo de manera independiente para cada muestra, debido a la naturaleza contingente del clima organizacional. Cada empresa tiene su propia historia, procesos de producción, estructura, orientación, filosofía y otros factores situacionales que la hacen única. Sin embargo, la validación en diferentes muestras va consolidando y robusteciendo al instrumento de medición. Algunas "pinceladas" de los resultados en una muestra se han presentado a

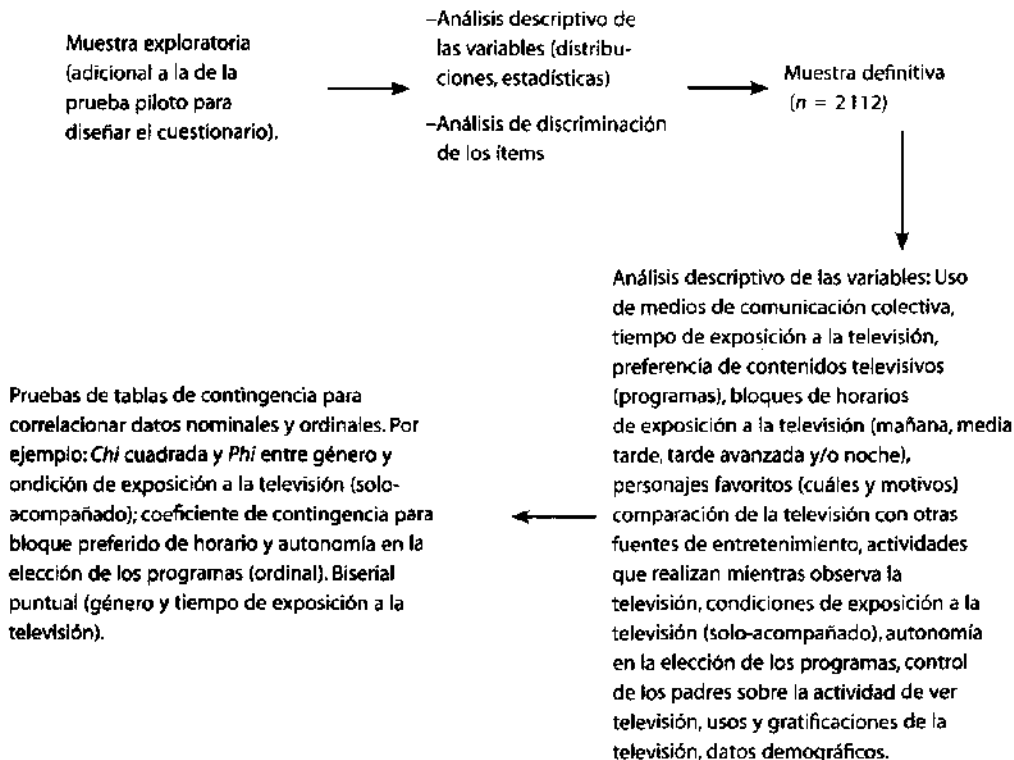


Figura 10.25 La secuencia de análisis con Minitab.

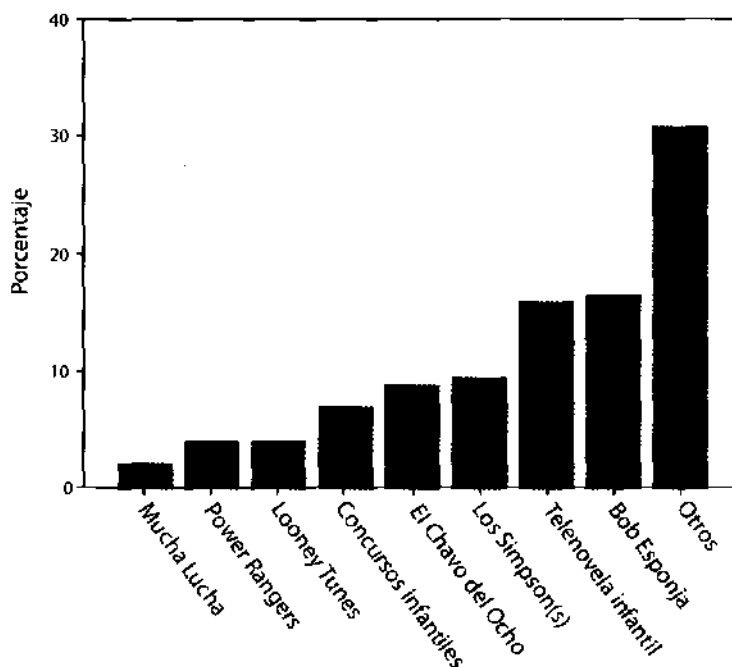


Figura 10.26 Programas preferidos (agrupados aquellos con menos de 4%).

lo largo del capítulo.⁹ Algunos de los resultados de dos de las muestras, se presentan en el capítulo 8 del CD anexo. Pero señalaremos el proceso de análisis y ciertos resultados de una de las muestras (institución educativa).

El programa elegido de análisis fue el SPSS. La secuencia de análisis fue la que se muestra en la figura 10.27.

La confiabilidad (*alfa*) del instrumento fue de 0.98 (muy considerable) ($n = 163$).¹⁰ El análisis de factores reveló una dimensión única y después de efectuarlo, se eliminaron tres ítems. Los coeficientes *alfa* para las dimensiones, fueron los que se muestran en la tabla 10.28:

Estadística descriptiva en el nivel de toda la escala.

En la tabla 10.29, se presentan las principales estadísticas de los resultados a toda la escala del clima organizacional.

El promedio (3.7) y la mediana (3.8) son bastante favorables e indican que esta organización posee un clima organizacional positivo. La respectiva normalización, se presenta en la gráfica de la figura 10.28.

Las correlaciones entre los componentes del clima organizacional para esta segunda muestra, se presentan en la tabla 10.30 (Matriz de correlaciones de Pearson). La mayoría se encuentra entre rangos que oscilan entre 0.55 y 0.69, es decir, correlaciones medias y considerables. Destacan las correlaciones entre comunicación y dirección (0.800), satisfacción y

⁹ Ver referencias a Hernández Sampieri (2005).

¹⁰ Recordemos que si no se responde a 100% de los ítems son excluidos del cálculo total los casos. La muestra total fue de 202 casos, que es el universo de la institución.

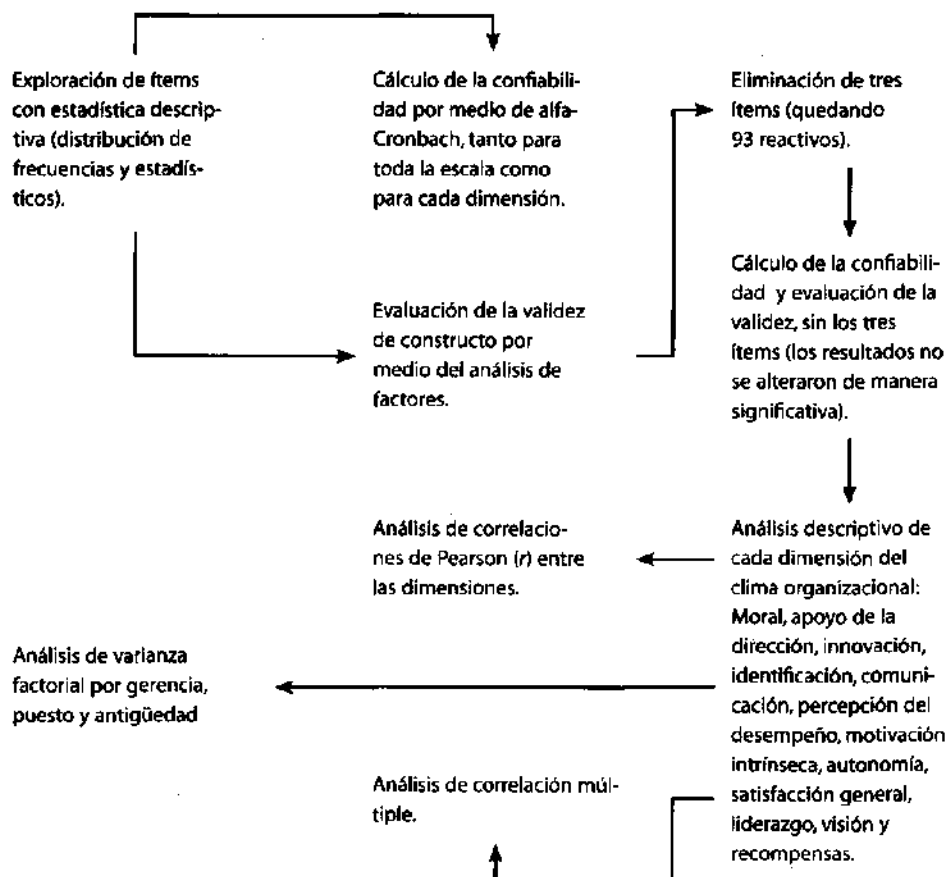


Figura 10.27 Secuencia de análisis con el SPSS.

Tabla 10.28 Los coeficientes *alfa* para las dimensiones del clima organizacional

Dimensión	n	Alfa
Moral	176	0.898
Dirección	176	0.928
Innovación	177	0.788
Identificación	175	0.823
Comunicación	179	0.827
Percepción del desempeño	174	0.709
Motivación intrínseca	178	0.792
Autonomía	179	0.827
Satisfacción	179	0.865
Liderazgo	177	0.947
Visión	176	0.902

Tabla 10.29 Estadística descriptiva de la muestra

Estadística	Valor
Media	3.7
Mediana	3.8
Moda	3.96
Desviación estándar	0.620.39
Varianza	-0.442
Asimetría	-0.116
Curtosis	1.84
Mínimo	4.92
Máximo	

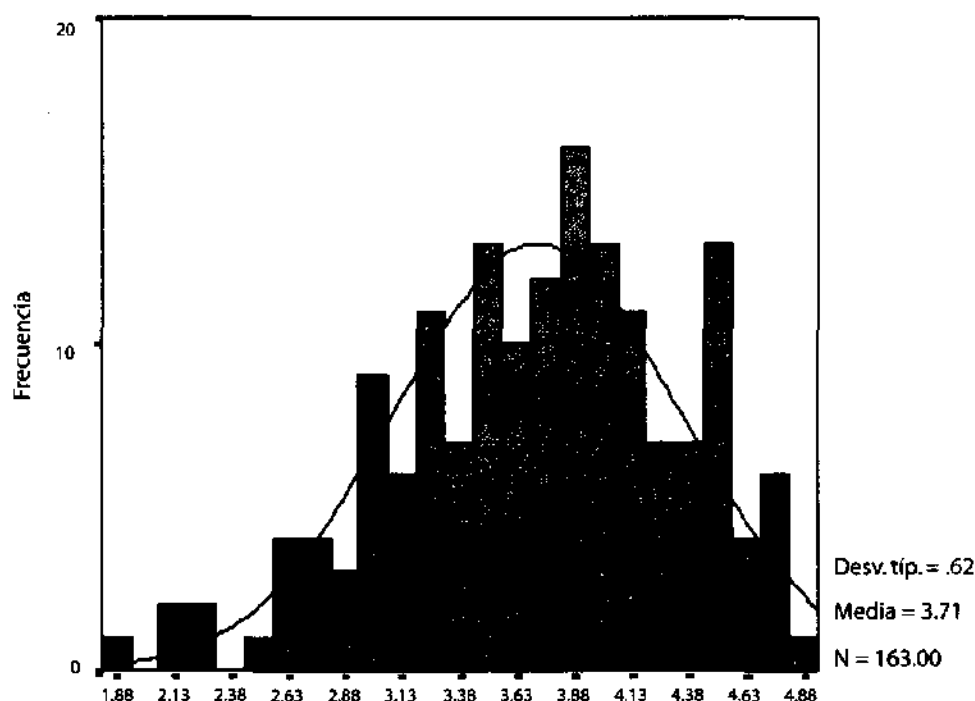


Figura 10.28 Histograma de toda la escala del clima organizacional en una institución educativa.

liderazgo (0.772), satisfacción y visión (0.721), identificación y motivación intrínseca (0.716); así como, moral y liderazgo (0.698). Llama la atención la correlación entre autonomía y motivación intrínseca

(0.613), que respalda la tradicional vinculación entre ambos conceptos, propuesta por los modelos sobre características del trabajo. Identificación y liderazgo es de las pocas correlaciones bajas (0.425).

Tabla 10.30 Matriz de correlaciones de Pearson

		MOR	DIR	INN	IDENT	COM	DES	MOT	AUT	SAT	LID	VIS
MOR	Correlación de	1	0.510*	0.684*	0.466*	0.583*	0.673*	0.588*	0.491*	0.681*	0.698*	0.692*
	Sig. (bilateral)	.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	176	174	174	173	175	171	176	176	175	176	173
DIR	Correlación de	0.510*	1	0.604*	0.601*	0.800*	0.588*	0.632*	0.649*	0.603*	0.589*	0.644*
	Sig. (bilateral)	0.000	.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	174	176	173	172	175	170	175	176	175	175	174
INN	Correlación de	0.684*	0.604*	1	0.544*	0.555*	0.612*	0.653*	0.505*	0.636*	0.640*	0.661*
	Sig. (bilateral)	0.000	0.000	.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	174	173	177	173	177	172	176	176	176	174	173
IDENT	Correlación de	0.466*	0.601*	0.544*	1	0.606*	0.556*	0.716*	0.539*	0.616*	0.425*	0.625*
	Sig. (bilateral)	0.000	0.000	0.000	.	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	173	172	173	175	174	170	175	174	174	173	171
COM	Correlación de	0.583*	0.800*	0.555*	0.606*	1	0.579*	0.550*	0.684*	0.659*	0.637*	0.647*
	Sig. (bilateral)	0.000	0.000	0.000	0.000	.	0.000	0.000	0.000	0.000	0.000	0.000
	N	175	175	177	174	179	173	177	178	178	176	175
DES	Correlación de	0.673*	0.588*	0.612*	0.556*	0.579*	1	0.682*	0.466*	0.579*	0.470*	0.672*
	Sig. (bilateral)	0.000	0.000	0.000	0.000	0.000	.	0.000	0.000	0.000	0.000	0.000
	N	171	170	172	170	173	174	173	173	174	171	171
MOT	Correlación de	0.588*	0.632*	0.653*	0.716*	0.550*	0.682*	1	0.613*	0.586*	0.532*	0.642*
	Sig. (bilateral)	0.000	0.000	0.000	0.000	0.000	0.000	.	0.000	0.000	0.000	0.000
	N	176	175	176	175	177	173	178	177	177	176	174
AUT	Correlación de	0.491*	0.649*	0.505*	0.539*	0.684*	0.466*	0.613*	1	0.616*	0.611*	0.585*
	Sig. (bilateral)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.	0.000	0.000	0.000
	N	176	176	176	174	178	173	177	179	178	177	176
SAT	Correlación de	0.681*	0.603*	0.636*	0.616*	0.659*	0.579*	0.586*	0.616*	1	0.772*	0.721*
	Sig. (bilateral)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.	0.000	0.000
	N	175	175	176	174	178	174	177	178	179	176	175
LID	Correlación de	0.698*	0.589*	0.640*	0.425*	0.637*	0.470*	0.532*	0.611*	0.772*	1	0.675*
	Sig. (bilateral)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.	0.000
	N	176	175	174	173	176	171	176	177	176	177	174
VIS	Correlación de	0.692*	0.644*	0.661*	0.625*	0.647*	0.672*	0.642*	0.585*	0.721*	0.675*	1
	Sig. (bilateral)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.
	N	173	174	173	171	175	171	174	176	175	174	176

*La correlación es significativa al nivel 0.01 (bilateral).

EL ABUSO SEXUAL INFANTIL

La confiabilidad de los instrumentos se resume en la tabla 10.31 (Confiabilidad de Instrumentos), para CKAQ-Español ($n = 150$, $\bar{x} = 5.08$, $DS = 3.43$, y rango de 8 a 22 puntos) y RP-México ($n = 150$, $\bar{x} = 11.53$, $DS = 7.97$ y un rango de 0 a 38 puntos).

Los tres grupos experimentales mostraron que existe un tipo de sensibilidad al instrumento en el CKAK-Español (Kruskal-Wallis $\chi^2 = 78.4$, $gl = 2$, $p < 0.001$) y RP-México (Kruskal-Wallis $\chi^2 = 83.06$, $gl = 2$, $p < 0.001$), lo cual indica que los grupos difieren en localización o forma, y reafirma la sensibilidad de las escalas al

Tabla 10.31 Confiabilidad de instrumentos

Instrumento	Confiabilidad interna ($p < 0.01$)	Confiabilidad de estabilidad temporal "test-retest" ($p < 0.01$)	Tipo de instrumento
CKAQ-Español	0.69	0.50	Cognitivo
RP-México	0.75	0.75	Conductual
EPA	0.78	0.75	Conductual

mostrar un comportamiento diferente entre los grupos, donde quienes más recientemente terminan el PPASI mejores puntajes obtienen.

Con el objetivo de indagar el comportamiento de los grupos de seguimiento y control con respecto al grupo que termina un PPASI, se calcularon los porcentajes relativos en las subescalas de hacer, decir, denunciar y el reconocimiento de los contactos positivos y negativos. Los resultados se exhiben en la tabla 10.32, que presenta los porcentajes de aciertos en relación con el grupo que acaba de concluir un PPASI en el RP-México. Se deduce que para las subescalas de reconocimiento de contactos negativos, DECIR y DENUNCIAR, se conserva el cambio esperado, quienes más recientemente hayan participado en un PPASI obtienen un mejor puntaje. En la habilidad de HACER, se observa que el grupo de seguimiento tiene

un mejor desempeño en promedio. Esto se puede explicar debido al incremento en la madurez de las niñas y niños, lograda a lo largo de un año aproximadamente; de 5.58 años en el primer grupo y de 6.47 años promedio en el grupo de seguimiento.

En la subescala de contactos positivos se advierte que al terminar el PPASI "Porque me quiero me cuido", tiene un puntaje promedio ligeramente menor que el grupo de control (11.45%), y mucho mejor en el grupo de seguimiento (53.71%), lo que avala que la escala es sensible a medir dicha habilidad y su posible efecto nocivo ante un PPASI. Este resultado si bien acredita parcialmente los resultados de Underwager y Wakefield (1993), que sostiene que al atender a los PPASI, los niños y niñas se muestran desconfiados(as) ante las aproximaciones cotidianas normales. También se constata que al cabo de un

Tabla 10.32 Porcentaje de rangos relativos con respecto al grupo que termina un PPASI

Grupo	Contactos negativos (%)	Contactos positivos (%)	Decir	Hacer	Denunciar
Al Terminar PPASI	100.00	100.00	100.00	100.00	100.00
Seguimiento	77.82	53.71	90.86	105.71	74.28
Control	37.39	115.45	44.14	49.14	41.52

año de concluido el programa, los infantes son capaces de superarlo y se muestran mucho más asertivos. Entre los grupos al terminar PPASI y de control se evidenció que solamente la habilidad de reconocer contactos positivos (Mann-Whitney $z = -1.48, n = 124, p = 0.14$) es la única habilidad que tiene la misma localización, esta evidencia contradice la teoría de Underwager y Wakefield (1993). Se puede concluir en este estudio que, si bien es cierto que al terminar el PPASI las niñas y niños aumentaban ligeramente el recelo ante los contactos positivos, esto no es significativo y con el tiempo, al incremento de la madurez, el fenómeno se supera.

En el caso de los grupos al terminar un PPASI y de seguimiento, se encontró que hay un mismo comportamiento en las subescalas de DECIR (Mann-Whitney $z = -1.20, n = 72, p = 0.23$) y HACER (Mann-Whitney $z = -1.26, n = 72, p = 0.21$) por lo que hay permanencia en el tiempo de estas dos habilidades.

La correlación entre las diferentes escalas verifica que hay una vinculación moderada entre las escalas CKAQ-Español y RP-México (Spearman $r = 0.68, n = 150, p < 0.01$), para el total de casos experimentales. En el caso de los grupos de control (Spearman $r = 0.23, n = 79, p < 0.05$) y al terminar el PPASI (Spearman $r = 0.35, p < 0.05$) las escalas tienen un nivel de correlación aún menor.

El resumen de puntajes por escala y grupo experimental se presenta en la tabla 10.33, donde la columna CKAQ-Español (Transformado) presenta la conversión de una escala con un puntaje de 0 a 22 a una de 0 a 40, para tener un comparativo equivalente entre las escalas cognitiva y conductual. Se observa que la escala cognitiva se encuentra por encima de la conductual en todos los grupos. Sin embargo, el grupo de control observa una mayor diferencia entre las escalas conductuales y cognitivas. El porcentaje relativo, con respecto al puntaje promedio del grupo al terminar PPASI, proporcional-

Tabla 10.33 Resumen descriptivo de puntajes por escala y grupo experimental

Grupo experimental	CKAQ-Español	CKAQ-Español (Transformado)	RP-México
Al terminar PPASI			
Media	18.33	33.32	19.11
Desviación estándar	2.66	4.83	6.48
Seguimiento			
Media	15.96	29.01	13.54
Desviación estándar	2.34	4.26	5.23
Control			
Media	12.81	23.29	5.96
Desviación estándar	2.17	3.94	4.37
Total			
Media	15.08	27.42	11.53
Desviación estándar	3.43	2.24	7.97

mente el grupo de seguimiento obtiene 87.07% de aciertos y el grupo de control 69.88%, para la escala CKAQ-Español. En el caso del instrumento conductual RP-México, el porcentaje relativo es de 70.85% de aciertos en el grupo de seguimiento y 31.19% en el grupo de control. Lo que evidencia que la sensibilidad al cambio en esta escala conductual es mayor.

Se observa también, que la distribución de la escala cognitiva CKAQ-Español, en general, se encuentra por encima de las escalas conductuales, lo que permite deducir que los menores pueden tener cierto grado de conocimiento que no se traduce en habilidades autoprotectoras.



Desde 1990 han disminuido las tensiones entre lo cualitativo versus lo cuantitativo, por lo que se buscó establecer una sinergia, así como ser más flexibles y eclécticos, dicho en el buen sentido, en los procedimientos.

La investigación cuantitativa ganó cuando particularizó los instrumentos y tomó en cuenta las características de los grupos a los cuales se dirige el estudio. Lo anterior propició un importante avance en la explicación de los procesos psicológicos, en especial los cognoscitivos; y en los descubrimientos neuropsicológicos, así como en el uso de software para el montaje de experimentos, demostraciones y simulaciones.

En este tipo de investigaciones, destacan las pruebas estadísticas por su utilidad en el análisis de datos categóricos de correspondencia, la ordenación de datos para conocer preferencias, el análisis factorial confirmativo, las correctas estimaciones de conjuntos de datos complejos, el manejo de resultados estadísticos de los experimentos, la validación de datos, la determinación del tamaño de la muestra y el análisis de regresión, entre otros aspectos a considerar.

A pesar de tan importantes avances en la investigación, aún hace falta financiamiento para una promoción significativa

y que, además, fomente la especialización de los investigadores, lo cual les permitiría competir de manera efectiva.

Ciro Hernando León Pardo

Coordinador del Área de Investigación

Facultad de Psicología

Universidad Javeriana

Bogotá, Colombia

Para efectuar una buena investigación se requiere plantear de forma correcta el problema, con lo cual tenemos 50 por ciento de la solución, y también con un rigor metodológico, es decir, incluir todos los pasos del proceso.

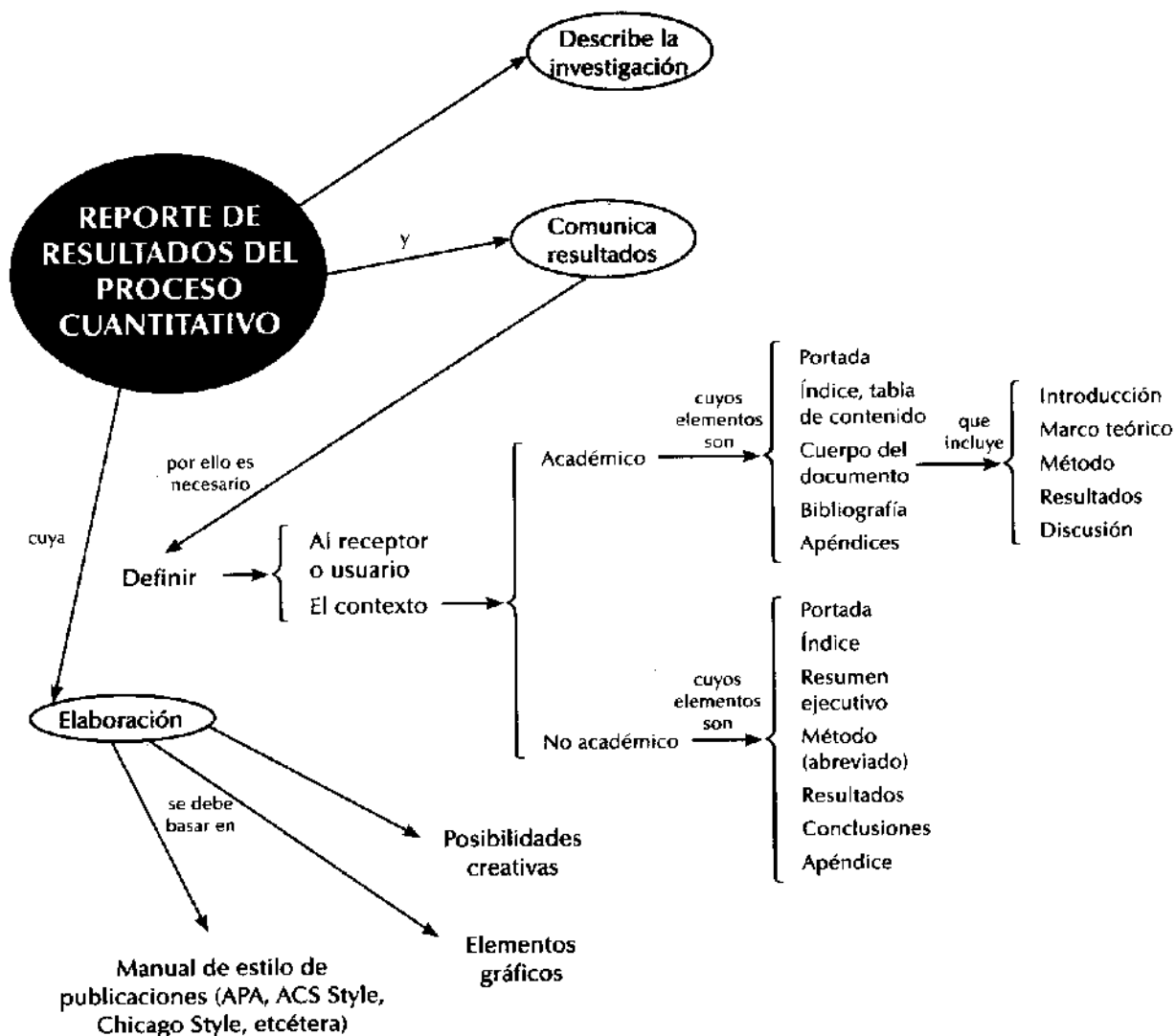
Tal apego a la metodología implica el empleo de los recursos pertinentes; por ejemplo, en las investigaciones sociales las pruebas estadísticas proporcionan una visión más precisa del objeto de estudio, ya que apoyan o no la hipótesis para su validación o rechazo.

Los estudiantes pueden concebir una idea de investigación a partir de sus intereses personales, aunque se recomienda que elijan temas íntimamente relacionados con su carrera, y que procuren que sean de actualidad y de interés común.

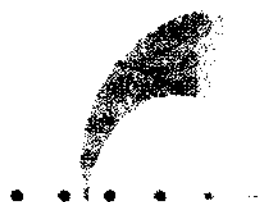
Para ello, los profesores deben infundir en los alumnos la importancia de la investigación en el terreno académico y en el

profesional, destacando su relevancia tanto en la generación de conocimiento como en la búsqueda de soluciones a problemas.

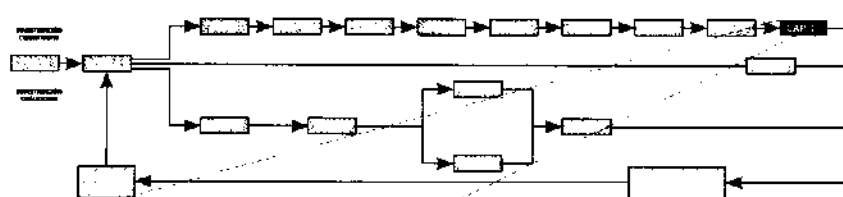
Roberto de Jesús Cruz Castillo
Profesor de tiempo completo
Facultad de Ciencias de la Administración
Universidad Autónoma de Chiapas
Chiapas, México



Capítulo 11



El reporte de resultados del proceso cuantitativo



PROCESO DE INVESTIGACIÓN CUANTITATIVA

Paso 10

Elaborar el reporte de resultados

- Definición del usuario.
- Selección del tipo de reporte a presentar: formato y contexto académico o no académico, dependiendo del usuario.
- Elaboración del reporte y del material adicional correspondiente.
- Presentación del reporte.

OBJETIVOS DE APRENDIZAJE

Al terminar este capítulo, el alumno será capaz de:

- Entender el papel tan importante que juega el usuario de la investigación en la elaboración del reporte de resultados.
- Reconocer los tipos de reportes de resultados en la investigación cuantitativa.
- Comprender los elementos que integran un reporte de investigación cuantitativa.

Síntesis

En el capítulo se comenta la importancia que tiene el usuario en la presentación de resultados. Éste es quien toma decisiones con base en los resultados de la investigación; por ello, la presentación debe adaptarse a sus características y necesidades.

Se mencionan dos tipos de reportes: académicos y no académicos, así como los elementos más comunes que integran un reporte producto de la investigación cuantitativa.

--- Antes de elaborar el reporte de investigación, debemos definir al receptor o usuario y al contexto

Se ha llevado a cabo una investigación y los resultados del estudio se han revisado, organizado e incluso, interpretado y comentado (se encuentran en tablas, gráficas, cuadros, diagramas, etc.); pero el proceso aún no termina. Es necesario *comunicar los resultados mediante un reporte*, el cual puede adquirir diferentes formatos: un libro, un artículo para una revista científica o para un diario de divulgación general, una presentación en computadora, un documento técnico, una tesis, un DVD, etc. En cualquier caso, debemos describir la investigación.

Usuario Persona que toma decisiones con base en los resultados de la investigación; por ello, la presentación debe adaptarse a sus necesidades.

Lo primero entonces es definir el tipo de reporte que es necesario elaborar, esto depende de varias precisiones: 1) las razones por las cuales surgió la investigación, 2) los usuarios del estudio y 3) el contexto en el cual se habrá de presentar. Por lo tanto, es necesario que antes de comenzar a desarrollar el reporte, el investigador reflexione respecto de las siguientes preguntas: ¿Cuál fue el motivo o los motivos que originaron el estudio? (que nadie conoce mejor que el investigador), ¿cuál es el contexto en que habrán de presentarse los resultados?, ¿quiénes son los usuarios de los resultados? y ¿cuáles son las características de tales usuarios? La manera en que se presenten los resultados dependerá de las respuestas a dichas preguntas.

Si el motivo fue elaborar una tesis para obtener un grado académico, el panorama es claro: el formato del reporte debe ser, justamente, una tesis de acuerdo con el grado que se cursó (licenciatura, maestría o doctorado) y los lineamientos a seguir son los establecidos por la institución educativa donde se habrá de presentar, el contexto será académico y los usuarios serán en primera instancia, los sinodales o miembros de un jurado y, posteriormente, otros alumnos y profesores de la propia universidad y otras organizaciones educativas. Si se trata de un trabajo solicitado por un profesor para una asignatura, materia o curso, el formato es un reporte académico cuyo usuario principal es el maestro que encargó el trabajo y los usuarios inmediatos son los compañeros que cursan la misma asignatura, para que después se agreguen como usuarios otros estudiantes de la escuela o facultad de nuestra institución y de otras universidades. En caso de que la razón que originó el estudio fue la solicitud de una empresa para que se analizara determinado aspecto que interesa a sus directivos. El reporte será en un contexto no académico y los usuarios básicamente son un grupo de ejecutivos de la empresa en cuestión que utilizará los datos para tomar ciertas decisiones. O en ocasiones, la investigación tiene varios motivos por los que se efectuó y diferentes usuarios (imaginemos que realizamos un estudio pensando en diversos productos y usuarios: un artículo que se someterá a consideración para ser publicado en una revista científica, una ponencia para ser presentada en un congreso, un libro, etc.). En este caso, suele primero elaborarse un documento central para después, desprender de éste distintos subproductos.

Vamos primero a considerar a los usuarios de la investigación, los contextos en que puede llevarse a cabo, los estándares que regularmente se contemplan al elaborar el reporte y que debemos tomar en cuenta, así como el tipo de reporte que comúnmente se utiliza en cada caso; los cuales se resumen en la tabla 11.1.

Los estándares son las bases para elaborar el reporte. La regulación en el campo académico casi siempre es mayor que en contextos no académicos, en los cuales no hay tantas reglas generales.

Tabla 11.1 Usuarios, contextos y estándares para la investigación¹

Usuarios	Contextos comunes posibles	Estándares que normalmente aplican para elaborar el reporte	Tipo de reporte
Académicos de la propia institución educativa: profesores, asesores, miembros de comités y jurados, alumnos. (Tesis y disertaciones, estudios institucionales para sus propias publicaciones o de interés para la comunidad universitaria.)	Académico	• Lineamientos utilizados en el pasado para regular las investigaciones en la escuela o facultad (o a nivel institucional). Es común que haya un manual institucional. • Lineamientos individuales de los decanos y profesores-investigadores de la escuela, facultad o departamento.	• Tesis y disertaciones • Informes de investigación • Presentaciones audiovisuales (Power Point, Flash, Dreamweaver, Slim Show, etcétera) • Libro
Editores y revisores de revistas científicas (<i>journals</i>)	Académico	• Lineamientos publicados por el editor y/o comité editorial de la revista (en ocasiones se diferencian por su tipo: si son investigaciones cuantitativas, cualitativas o mixtas). Es común que se denominen "normas o instrucciones para los autores". El tema de nuestro estudio debe encuadrar dentro del tema de la revista y a veces en el volumen en cuestión (que puede ser anual o bianual).	Artículos
Revisores de ponencias para congresos y académicos externos (ponencias, presentaciones en congresos, foros en internet, páginas web, premios a la investigación, etcétera).	Académico	• Lineamientos o estándares definidos en la convocatoria del congreso, foro o certamen. Estos estándares son para el escrito que se presenta y/o publica, así como para los materiales adicionales requeridos (por ejemplo, presentación visual, video, resumen gráfico para cartel). El tema de nuestro estudio debe encuadrar dentro del tema de la conferencia y tenemos que ajustarnos a la normatividad definida para las ponencias.	• Ponencias • Póster o cartel

(continúa)

Tabla 11.1 Usuarios, contextos y estándares para la investigación (*continuación*)

Usuarios	Contextos comunes posibles	Estándares que normalmente aplican para elaborar el reporte	Tipo de reporte
Elaboradores de políticas, ejecutivos o funcionarios que toman decisiones (empresas, organizaciones gubernamentales y organizaciones no gubernamentales).	Académico No académico (regularmente el caso de las empresas)	• Lineamientos lógicos o estándares utilitarios: – Informe breve, cuyos resultados sean fáciles de entender. – Orientación más bien visual del contenido (gráficas, cuadros, etc.; solamente los elementos más importantes). – Posibilidad de aplicar los resultados de manera inmediata. – Claridad de ideas.	• Resumen ejecutivo • Informe técnico • Presentaciones audiovisuales
Profesionistas y practicantes dentro del campo donde se inserta el estudio.	Académico No académico (regularmente el caso de las empresas)	• Lineamientos lógicos o estándares pragmáticos: – Relevancia del problema estudiado. – Orientación más bien visual del contenido (gráficas, cuadros, etc.; sólo los elementos más importantes). – Resultados fácilmente identificables y aplicables. – Sugerencias prácticas y concretas para implementar.	• Resumen ejecutivo • Informe técnico • Presentaciones audiovisuales
Opinión pública no especializada (estudiantes de primeros ciclos, padres de familia, grupos de la sociedad en general).	No académico	• Estándares centrados en la sencillez de los resultados, su importancia para un grupo de la sociedad o ésta en su conjunto: – Brevedad. – Claridad. – Aplicabilidad a situaciones cotidianas. – Orientación más bien visual del contenido (gráficas, cuadros, etc.; pocos elementos, dos o tres muy sencillos).	• Artículo periodístico • Libro

¹ Adaptado de Creswell (2005, p. 258).

Las tesis y disertaciones varían en extensión, pues éstas dependen del estudio en sí y las normas institucionales. Aunque la tendencia actual es incluir los elementos y contenidos realmente necesarios (en el pasado se exigían cientos de páginas y varias eran desechables).

Los informes y libros también varían en cuanto a su tamaño y, asimismo, éste depende de la investigación en sí y el detalle con que los usuarios de la investigación deseen los datos. Los artículos para revistas científicas rara vez son mayores de 30 páginas.² Los pósters o carteles normalmente son de una o dos páginas de acuerdo al tamaño que sea requerido por los organizadores de la conferencia. Los escritos para presentarse como ponencias suelen no exceder de 30 minutos, pero también depende del comité que organiza cada acto académico. Los artículos periodísticos regularmente no ocupan más de una página del diario, en el caso más extenso.

¿Qué elementos contiene un reporte de investigación o un reporte de resultados en un contexto académico?

Los elementos que contiene la mayoría de los reportes dependen del tipo de reporte y usuarios o lectores potenciales de la investigación. Sin embargo, varios autores³ sugieren los elementos que a continuación veremos completos.

El reporte Siempre se compara con la propuesta, proyecto o protocolo (este tema se incluye en el CD anexo).

1. **Portada.** Incluye el título de la investigación; el nombre del autor o los autores y su afiliación institucional, o el nombre de la organización que patrocina el estudio, así como la fecha y el lugar en que se presenta el reporte.
En el caso de tesis y disertaciones, las portadas varían de acuerdo con los lineamientos establecidos por la autoridad pública o la institución de educación superior correspondiente.
2. **Índice del reporte o tabla de contenidos.** Incluye capítulos, apartados y subapartados (numerados o diferenciados por tamaño y características de la tipografía). Frecuentemente se agregan índices de tablas y/o figuras, cuadros y diagramas.
3. **Resumen.** Que constituye brevemente el contenido esencial del reporte de investigación, y usualmente incluye el planteamiento del problema, el método, los resultados más importantes y las principales conclusiones. Debe ser comprensible, sencillo, informativo, preciso, completo, conciso y específico (American Psychological Association, 2002). En el caso de artículos para revistas científicas, no se recomienda exceder las 120 palabras (American Psychological Association, 2002). En tesis, se sugiere que no exceda las 320 palabras (el estándar es de 300). En el recuadro de ejemplos se presentan modelos de resúmenes de un artículo de revis-

² Williams, Tutty y Grinnell (2005); *Journal of Communication* (2005); *Quarterly Journal of Business and Economics* (2005); *Archivos Hispánicos de Sexología* (2005); *International Journal of Clinical and Health Psychology* (2005) que es una revista editada conjuntamente por la Asociación Española de Psicología Conductual, la Asociación Colombiana para el Avance de las Ciencias del Comportamiento y la Associação Portuguesa de Terapia do Comportamento. Por ejemplo, el *European Journal on Human Genetics* (2005) solicita que los artículos no sean mayores de 20 páginas y el *European Journal of Information Systems* (2005) solamente 10 páginas o seis mil palabras. Por su parte, el *International Journal of Hindu Studies* (2005) tiene una limitante de 40 páginas, mientras que *Social Science Japan Journal* (2005) circunscribe los artículos a 10 mil palabras (alrededor de 17 páginas).

³ Se tomaron elementos de Creswell (2005), Williams, Tutty y Grinnell (2005), Hernández Sampieri, Fernández Collado y Baptista (2003), American Psychological Association (2002) y Laflen (2001).

ta científica, una investigación, una tesis y una ponencia. Casi en todas las revistas académicas y tesis se exige que el resumen esté en el idioma original en el cual se produjo el estudio (en nuestro caso en español) y en inglés. El resumen regularmente abarca: a) el planteamiento en un solo párrafo, b) la muestra en otro párrafo (tamaño y características —por ejemplo, si son participantes, número, tipo, edad, género, origen étnico—), c) el método en otro párrafo (diseño, instrumentos, procedimientos), d) resultados más importantes en máximo dos párrafos (prueba de hipótesis, incluyendo niveles de significancia o significación), e) la conclusión más relevante o central y sus implicaciones fundamentales (último párrafo).

EJEMPLOS

Resumen en inglés de un artículo de revista científica (Rubin, Fernández Collado y Hernández Sampieri, 2005)

A CROSS-CULTURAL EXAMINATION OF INTERPERSONAL COMMUNICATION MOTIVES IN MÉXICO AND THE UNITED STATES

REBECCA B. RUBIN

Kent State University

CARLOS FERNÁNDEZ COLLADO and

ROBERTO HERNÁNDEZ-SAMPIERI

Universidad Anáhuac

ABSTRACT. *This study examined cultural differences in motives for interpersonal communication by comparing a general U.S. sample with student samples from the U.S. and México. Past research indicated six main motives that people have for initiating conversations with others: Pleasure, escape, relaxation, inclusion, affection, and control. Hofstede's (1980) dimensions of national culture—power distance, uncertainty avoidance, individualism, and masculinity were used to predict cross-cultural differences in these interpersonal motives.*

Eight hypotheses were tested. Results indicated that Mexican scores were not significantly higher than U.S. scores on interpersonal control, relaxation, and escape motives (as predicted), but were significantly lower than U.S. scores on interpersonal affection, pleasure, and inclusion motives (as predicted). There were significant negative relationships between the interpersonal communication motives and age in the U.S. data, but not in the Mexican data. Also, significant positive relationships between gender and the affection and inclusion motives and negative relationship between gender and control motives were found only for the U.S. samples.

Resumen de una investigación que podría ser una tesis (Hernández Sampieri, 2005)⁴

Resumen

La investigación estableció como objetivo validar un instrumento para medir el clima organizacional. El procedimiento de validación se efectuó en dos muestras independientes: Un laboratorio químico-farmacéutico y una institución educativa ($n = 421$ y $n = 202$, respectivamente).

Se desarrolló el instrumento de medición sobre la base de la revisión de la literatura y de otros instrumentos validados previamente. El instrumento consistió en un cuestionario de 100 reactivos o ítems, el cual se redujo a 96 después de la prueba piloto. Las dimensiones del clima organizacional medidas fueron: moral, apoyo de la dirección, innovación, identificación, comunicación, percepción del desempeño, motivación intrínseca, autonomía, satisfacción general, liderazgo y visión. En el caso de la muestra del laboratorio fue agregada la dimensión "recompensas y compensaciones".

El análisis de reducción de datos generó 19 y 17 factores, respectivamente, pero en ambas muestras los ítems se desarrollaron con tendencia hacia un solo factor, lo que conduce a la conclusión de que el clima organizacional es un constructo "molar", en el cual "se funden" distintas percepciones sobre aspectos centrales del ambiente de trabajo. Los coeficientes de confiabilidad *alfa*-Cronbach resultantes fueron de 0.98 en las dos muestras.

La mayoría de las dimensiones del clima organizacional mostraron correlaciones significativas entre sí, tanto para el laboratorio como para la institución educativa.

En la primera muestra, el análisis de regresión reveló que todas las dimensiones, excepto recompensas, resultaron buenos predictores de la escala completa del clima organizacional (R cuadrada = 0.988). En la segunda muestra no se obtuvieron coeficientes *beta* significativos en: percepción del desempeño, motivación intrínseca, liderazgo y autonomía (R cuadrada = 0.799).

El instrumento fue validado y puede utilizarse en el medio laboral mexicano, tomando en cuenta las sugerencias que se comentan a lo largo de la tesis.

Descriptores: Clima organizacional, medición, instrumentos de medición

⁴ La investigación se ha presentado a lo largo del libro como ejemplo. A veces se considera una sola de las muestras del estudio, en otras se comenta la fase inicial del estudio que incluyó cuatro muestras. En este caso, se presentan un primer reporte que abarcó a las dos primeras muestras: laboratorio químico-farmacéutico e institución educativa mediana. El estudio concluirá a mediados de 2006 con 10 organizaciones y en 2007 con 20 en total. El estudio generó varios subproductos y se han citado partes de éstos: artículos, documentos técnicos, tesis de grado, etc.; por ello habla de distintas muestras. Se comenta hasta este capítulo para no desconcertar al lector.

Resumen de una ponencia (con apartados) de Meza y Hernández Sampieri (2005)

Comparativo de instrumentos de evaluación para programas de prevención del abuso sexual infantil en preescolares.

María Elena Meza de Luna y Roberto Hernández Sampieri

Universidad de Celaya

Resumen

Objetivo: Este estudio compara el comportamiento de dos medidas, una cognitiva y la otra conductual para evaluar los programas de prevención del abuso sexual en niñas y niños entre cuatro y seis años de edad.

Método: Investigación efectuada con preescolares ($n = 150$), en tres condiciones experimentales, quienes contestaron a un instrumento cognitivo (*Children's Knowledge Abuse Questionnaire*, versión en español) y a otro conductual (*Role Play Mexico*).

Resultados: A menores grados de formación de habilidades asertivas y auto—defensivas, menor es el grado de correlación entre las escalas conductuales y cognitivas. El instrumento conductual mostró una mejor sensibilidad ante las diferentes condiciones experimentales y una confiabilidad más elevada.

Conclusión: Este estudio aporta evidencia para recomendar más bien el empleo de modelos de evaluación conductuales que los cognitivos en preescolares. Cabe destacar que evaluar sobre la cognición de conceptos no garantiza una habilidad autoprotectora.

Descriptor: Instrumentos de Evaluación, Cognitivo, Conductual, Abuso Sexual, Prevención.

4. Cuerpo del documento.

- **Introducción.** Incluye los antecedentes (brevemente tratados de manera concreta y específica), el planteamiento del problema (objetivos y preguntas de investigación, así como la justificación del estudio), el contexto de la investigación (cómo y dónde se realizó), las variables y los términos de la investigación y sus definiciones conceptuales y operacionales, lo mismo que las limitaciones de ésta. Es importante que se comente la utilidad del estudio para el campo profesional. Creswell (2005) le denomina el planteamiento del problema y agrega las hipótesis. Laflén (2001) recomienda una serie de preguntas para elaborar la introducción: ¿qué descubrió o probó la investigación?, ¿en qué clase de problema se trabajó, cómo se trabajó y por qué se trabajó de cierta manera?, ¿qué motivó el estudio?, ¿por qué se escribe el reporte? y ¿qué debe saber o entender el lector al terminar de leer el reporte?
- **Marco teórico** (marco de referencia o revisión de la literatura). En el que se incluyen y comentan las teorías que se manejaron y los estudios previos que fueron relacionados con el planteamiento, se hace un sumario de los temas y hallazgos más importantes en el pasado, y se señala cómo el estudio amplió la literatura actual. Para mayor orientación sobre su desarrollo, se recomienda leer el cuarto capítulo del presente libro ("Elaboración del marco teórico"), pero

finalmente nos debe responder la pregunta: ¿Dónde estamos ubicados actualmente en cuanto al conocimiento referente a nuestras preguntas y objetivos?

- **Método.** Esta parte del reporte describe cómo fue llevada a cabo la investigación, e incluye:
 - Enfoque (cuantitativo, cualitativo o mixto).
 - Contexto de la investigación (lugar o sitio y tiempo, así como accesos y permisos).
 - Casos, universo y muestra (tipo, procedencia, edades, género o aquellas características que sean relevantes de los casos; descripción del universo y la muestra, y procedimiento de selección de la muestra).
 - Diseño utilizado (experimental o no experimental, así como intervenciones, si es que se utilizaron).
 - Procedimiento (un resumen de cada paso en el desarrollo de la investigación). Por ejemplo, en un experimento se describe la manera de asignar los participantes a los grupos, las instrucciones, los materiales, las manipulaciones experimentales, y cómo transcurrió el experimento. En una encuesta se describe cómo se contactó a los sujetos y se realizaron las entrevistas. En este rubro se incluyen los problemas enfrentados y la forma en la que se resolvieron.
 - Descripción detallada de los procesos de recolección de los datos y qué se hizo con los datos una vez obtenidos.
 - En cuanto a la recolección, es necesario describir qué datos fueron recabados, cuándo fueron recogidos y cómo: forma de recolección y/o instrumentos de medición utilizados, con *reporte de la confiabilidad, validez y objetividad, así como las variables o conceptos, eventos, situaciones y categorías.*
- **Resultados.** Éstos son producto del análisis de los datos. Compendian el tratamiento estadístico que se les dio a los datos. Regularmente el orden es: a) análisis descriptivos de los datos, b) análisis inferenciales para responder a las preguntas y/o probar hipótesis (en el mismo orden en que fueron formuladas las hipótesis o las variables). La American Psychological Association (2002) recomienda que primero se describa de manera breve la idea principal que resume los resultados o descubrimientos, y posteriormente se reporten con detalle los resultados. Es importante destacar que en este apartado no se incluyen conclusiones ni sugerencias, así como tampoco se explican las implicaciones de la investigación. Esto se hace en el siguiente apartado.

En el apartado de resultados, el investigador se limita a describir sus hallazgos. Una manera útil de hacerlo es mediante tablas, cuadros, gráficas, dibujos, diagramas, mapas y figuras generados por el análisis. Son elementos que sirven para organizar los datos, de tal manera que el usuario o lector los pueda leer y decir: “me queda claro que esto se vincula con aquello, con esta variable ocurre tal cuestión...” Cada uno de dichos elementos debe ir numerado (en arábigo o romano) (por ejemplo: cuadro 1, cuadro 2... cuadro *k*; gráfica o diagrama 1, gráfica o diagrama 2... gráfica o diagrama *k*, etc.) y con el título que lo identifica. Wiersma y Jurs (2005) recomiendan los siguientes puntos para elaborar tablas estadísticas:

- a) El *título* debe especificar el contenido de la tabla, así como tener un *encabezado* y los *subencabezados* necesarios (por ejemplo, columnas y renglones, diagonales, etcétera).
- b) No debe mezclarse una cantidad inmanejable de estadísticas (por ejemplo, incluir medias, desviaciones estándar, correlaciones, razón *F*, etc., en una misma tabla).

- c) En cada tabla se deben *espaciar los números y las estadísticas incluidas* (tienen que ser legibles).
- d) De ser posible, habrá que *limitar cada tabla a una sola página*.
- e) Los formatos de las tablas tienen que ser coherentes y homogéneos dentro del reporte (por ejemplo, no incluir en una tabla cruzada las categorías de la variable dependiente en columnas y en otra tabla colocar las categorías de la variable dependiente en renglones).
- f) Las *categorías de las variables deben distinguirse* claramente entre sí.

La mejor regla para elaborar una tabla es organizarla lógicamente y eliminar la información que pueda confundir al lector. Al incluir pruebas de significancia: *F*, *chi cuadrada*, *r*, etc., debe incorporarse información respecto de la magnitud o el valor obtenido de la prueba, los grados de libertad, el nivel de confianza (*alfa* = α) y la dirección del efecto (American Psychological Association, 2002). Asimismo, tendrá que especificarse si se acepta o se rechaza la hipótesis de investigación o la nula en cada caso. En el siguiente ejemplo se presenta un estudio realizado en tres plantas de una empresa para compararlas en su motivación intrínseca. El instrumento fue una escala tipo Likert con 10 ítems, por lo cual el rango posible era de 10 a 50 (la opción más negativa se calificó con uno y la más alta con cinco). La hipótesis de investigación es: Las plantas diferirán en cuanto a la motivación intrínseca (ver ejemplo).

EJEMPLO

Gráficas y tablas que pueden reducirse

- 1) Las estadísticas descriptivas para toda la muestra fueron las siguientes:

Motivación intrínseca en el trabajo

<i>N</i>	Válidos	306
	Perdidos	5
Media		30.33
Error típ. de la media		0.613
Mediana		30.00
Moda		22
Desv. típ.		10.725
Varianza		115.022
Asimetría		0.094
Error típ. de asimetría		0.139
Curtosis		-1.358
Error típ. de curtosis		0.278
Rango		37
Mínimo		12
Máximo		49
Suma		9 281

Resulta obvio que es demasiada información, sería mejor resumirla e incluir sólo las estadísticas más importantes:

Motivación intrínseca en el trabajo

<i>N</i>	Válidos Perdidos	306 5
Media		30.33
Mediana		30.00
Moda		22
Desv. típ.		10.725
Mínimo		12
Máximo		49

- 2) Y para la estadística descriptiva por planta, en lugar de una tabla larga con los valores de cada una, mejor se mencionan los promedios y desviación estándar de las tres: planta EC, planta Central y planta RH.

En lugar de la siguiente tabla:

Motivación intrínseca en el trabajo

					Intervalo de confianza para la media al 95%			
	<i>N</i>	Media	Desviación típica	Error típico	Límite inferior	Límite superior	Mínimo	Máximo
Planta EC	90	41.19	5.728	0.604	39.99	42.39	24	49
Planta Central	126	29.27	8.641	0.770	27.75	30.79	12	47
Planta RH	90	20.96	6.842	0.721	19.52	22.39	12	40
Total	306	30.33	10.725	0.613	29.12	31.54	12	49

Tendríamos el siguiente comentario: El promedio y desviación estándar de las plantas fue: EC (41.2, 5.7), Central (29.3, 8.6) y RH (21, 6.8) o una tabla más pequeña:

	<i>N</i>	Media	Desviación típica
Planta EC	90	41.19	5.728
Planta Central	126	29.27	8.641
Planta RH	90	20.96	6.842
Total	306	30.33	10.725

Y en torno a la prueba de la hipótesis, se tendrían los resultados del análisis de varianza (comparaciones en motivación intrínseca entre las tres plantas). En lugar de esta tabla:

Motivación intrínseca en el trabajo

	Suma de cuadrados	<i>gl</i>	Media cuadrática	<i>F</i>	Sig.
Intergrupos	18 663.227	2	9 331.613	172.214	0.000
Intragrupos	16 418.437	303	54.186		
Total	35 081.663	305			

Es mejor comentar: Las tres plantas difirieron significativamente en lo relativo a la motivación intrínseca ($F = 172.2, s < 0.01, gl = 2$), se acepta la hipótesis de investigación. Los empleados de la planta EC poseen altos niveles de motivación intrínseca, los empleados de la planta central, valores medios; y los de la planta RH, bajos. O al menos, se eliminan de la tabla las columnas: fuente de variación inter e intragrupos, suma de cuadrados y media cuadrática.

Cuando los *usuarios*, receptores o lectores son personas con conocimientos sobre estadística no es necesario explicar en qué consiste cada prueba, sólo habrá que mencionarlás y comentar sus resultados (que es lo normal en ambientes académicos). Si el usuario carece de tales conocimientos, no tiene caso incluir las pruebas estadísticas, a menos que se expliquen con suma sencillez y se presenten los resultados más comprensibles. En este caso, las tablas se describen.

En el caso de diagramas, figuras, mapas cognoscitivos, esquemas, matrices y otros elementos gráficos, también debe seguirse una secuencia de numeración y observar el principio básico: *una buena figura es sencilla, clara y no estorba la continuidad de la lectura*. Las tablas, los cuadros, las figuras y los gráficos tendrán que enriquecer el texto; en lugar de duplicarlo, comunican los hechos esenciales, son fáciles de leer y comprender, a la vez que son coherentes.

- *Conclusiones, recomendaciones e implicaciones* (o discusión). En esta parte se: a) derivan conclusiones, b) explicitan recomendaciones para otras investigaciones (por ejemplo, sugerir nuevas preguntas, muestras, instrumentos) y se indica lo que sigue y lo que debe hacerse, c) generalizan los resultados a la población, d) evalúan las implicaciones de la investigación, e) establecen cómo se respondieron las preguntas de investigación, así como si se cumplieron o no los objetivos, f) relacionan los resultados con los estudios existentes (vincular con el marco teórico y señalar si nuestros resultados coinciden o no con la literatura previa, en qué sí y en qué no), g) comentan las limitaciones de la investigación, h) destaca la importancia y significado de todo el estudio, i) explican los resultados inesperados y j) cuando no se probaron las hipótesis es necesario señalar o al menos especular sobre las razones. Al elaborar las conclusiones es aconsejable verificar que estén los puntos necesarios, aquí vertidos. Y no olvidar que no se trata de repetir los resultados, sino de resumir los resultados más importantes. Desde luego, las conclusiones deben ser congruentes con los datos. La adecuación de las conclusiones respecto de la generalización de los resultados deberá evaluarse en términos de aplicabilidad a diferentes muestras y poblaciones. Si el planteamiento cambió, es necesario explicar por qué y cómo se modificó. Esta parte debe redactarse de tal manera que se facilite la toma de decisiones respecto de una teoría, un curso de acción o una problemática. El reporte de un experimento tiene que explicar con claridad las influencias de los tratamientos.
5. **Referencias o bibliografía.** Son las fuentes primarias utilizadas por el investigador para elaborar el marco teórico u otros propósitos; se incluyen al final del reporte, ordenadas alfabéticamente. Cuando un mismo autor aparezca dos veces, debemos organizar las referencias que lo contienen de la más antigua a la más reciente.
 6. **Apéndices.** Resultan útiles para describir con mayor profundidad ciertos materiales, sin distraer la lectura del texto principal del reporte, o evitar que dichos materiales rompan con el formato del reporte. Algunos ejemplos de apéndices serían el cuestionario utilizado, un nuevo programa computacional, análisis estadísticos adicionales, el desarrollo de una fórmula complicada, fotografías, etcétera.

Cabe destacar que en reportes para publicarse como los artículos de una revista científica, se excluyen la introducción y el resto de los elementos se desarrollan de manera muy concisa o resumida, y rara vez se incluyen apéndices. En todo momento debe buscarse claridad, precisión y explicaciones directas, así como eliminar repeticiones, argumentos innecesarios y redundancia no justificadas. En el lenguaje debemos ser muy cuidadosos y sensibles, no debemos utilizar términos despectivos refiriéndonos a personas con capacidades distintas, grupos étnicos diferentes al nuestro, etc.; para ello, es necesario consultar algún manual de los que se recomiendan más adelante.

¿Qué elementos contiene un reporte de investigación o reporte de resultados en un contexto no académico?

Un reporte no académico contiene la mayoría de los elementos de un reporte académico:

1. Portada.
2. Índice.
3. Resumen ejecutivo (resultados más relevantes y casi todos presentados de manera gráfica).
4. Método.
5. Resultados.
6. Conclusiones.
7. Apéndices.

Pero cada elemento se trata con mayor brevedad y se eliminan las explicaciones técnicas que no puedan ser comprendidas por los usuarios. El marco teórico y la bibliografía suelen omitirse del reporte o se agregan como apéndices o antecedentes. Desde luego, lo anterior de ninguna manera implica que no se haya desarrollado un marco teórico, sino que algunos usuarios prefieren no confrontarse con éste en el reporte de investigación. Hay usuarios no académicos que sí se interesan por el marco teórico y las citas bibliográficas o referencias. Para ilustrar la diferencia entre redactar un reporte académico y uno no académico, se presenta un ejemplo de introducción de un reporte no académico que, como se ve en el siguiente recuadro de ejemplo, es bastante sencillo, breve, y no utiliza términos complejos.

EJEMPLO

Muestra de introducción de un reporte no académico

Calidad total

La Fundación Mexicana para la Calidad Total, A.C. (Fundameca), realizó una investigación por encuestas para conocer las prácticas, técnicas, estructuras, procesos y temáticas existentes en calidad total en nuestro país. La investigación, de carácter exploratorio, constituye el primer esfuerzo para obtener una radiografía del estado de los procesos de calidad en México. No es un estudio exhaustivo, sólo implica un primer acercamiento, que en los años venideros irá extendiendo y profundizando la Fundación.

El reporte de investigación que a continuación se presenta tiene como uno de sus objetivos esenciales propiciar el análisis, la discusión y la reflexión profunda respecto de los proyectos para incrementar la calidad de los productos o servicios que ofrece México a los mercados nacional e internacional. Como nación, sector y empresa: ¿Vamos por el camino correcto hacia el logro de la calidad total? ¿Qué estamos haciendo adecuadamente? ¿Qué nos falta? ¿Cuáles son los obstáculos a los que nos estamos enfrentando? ¿Cuáles son los retos que habremos de rebasar en la primera década del siglo? Éstas son algunas de las preguntas que estamos valorando y necesitamos responder. La investigación pretende aportar algunas pautas para que comencemos a contestar en forma satisfactoria dichos cuestionamientos.

La muestra de la investigación fue seleccionada al azar sobre la base de tres listados: listado *Expansión 500*, listado de la gaceta *Cambio Organizacional* y listado de las reuniones para constituir Fundameca. Se acudió a 184 empresas, de las cuales 60 no proporcionaron información. Dos encuestas fueron eliminadas por detectarse inconsistencias lógicas. En total se incluyeron 122 casos válidos.

Esperamos que sus comentarios y sugerencias amplíen y enriquezcan este proceso investigativo.

Fundameca

Dirección de Investigación

¿Dónde podemos consultar los detalles relativos a un reporte de investigación? (guías)

En la actualidad hay varios manuales que pueden ser útiles para elaborar los reportes:

1. *Manual de estilo de publicaciones* de la American Psychological Association (APA). Cubre todo lo relativo a cómo presentar un reporte de investigación. Se publica en inglés y en español. Vale la pena adquirirlo en la librería de su preferencia, es sumamente completo (abarca desde cómo citar hasta detalles de tablas y referencias). Además del manual hay un programa: APA-Style Helper.

Para mantenerse actualizado en nuevas ediciones del manual en inglés, la American Psychological Association tiene un sitio en internet: (<http://www.apastyle.org>). Una página de ayuda sobre el estilo APA es: (<http://www.psywww.com/resource/apacrib.htm>). Asimismo, tiene una página guía para preparar manuscritos que se piensen publicar en revistas, la cual se descarga en un PDF (*Portable Document File*): (<http://www.apa.org/journals/authors/guide.html>).

2. *The ACS Style Guide: A Manual for Authors and Editors*, editado por Janet S. Dodd y publicado por The American Chemical Society y Oxford University Press. Dirigido más bien a reportes de investigadores en ciencias químicas.
3. *Requisitos uniformes para la entrega de los manuscritos a las revistas biomédicas: la escritura y la edición para la publicación biomédica*, que está basado en el documento oficial: *Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication*. Publicado por el Comité Internacional de Directores de Revistas Médicas (CIDRM). Disponible en español sin costo en: (http://bvs.sld.cu/revistas/aci/vol12_3_04/aci12304.htm).

El sitio del documento original en inglés es: (<http://www.icmje.org/>).

De los manuales, siempre es necesario buscar la edición más actualizada.

4. Las páginas de cada revista académica (*journals*) en la sección: instrucciones para autores, también son muy útiles, en lo referente a artículos.

¿Qué recursos están disponibles para presentar el reporte de investigación?

Son hoy tantos los programas de dibujo, de gráficas, presentaciones y elaboración de documentos, que es imposible en este espacio comentarlos o siquiera nombrarlos. Use todos los que conozca y tenga acceso a ellos, recuerde que una presentación debe tener riqueza visual. En los documentos hay ciertas reglas que no podemos hacer a un lado, pero en la presentación el límite es nuestra propia imaginación.

¿Qué criterios o parámetros podemos definir para evaluar una investigación o un reporte?

Una propuesta de parámetros o criterios para evaluar la calidad de un estudio y, consecuentemente, su reporte, se presenta en el capítulo 10 del CD anexo por cada etapa del proceso de investigación cuantitativa y es producto de las observaciones de varios autores y profesores en Iberoamérica. Es una manera de visualizar si hemos llevado a cabo una investigación de manera rigurosa.

¿Con qué se compara el reporte de la investigación?

El reporte se contrasta con la propuesta o protocolo de la investigación, la que hicimos al inicio del proceso, que no se ha comentado en el libro, porque primero resultaba necesario conocer el proceso de investigación cuantitativa. El protocolo se revisará en el capítulo 9 del CD anexo: "Elaboración de propuestas cuantitativas, cualitativas y mixtas".

- Antes de elaborar el reporte de investigación debe definirse al usuario, ya que el reporte habrá de adaptarse a éste.
- Los reportes de investigación pueden presentarse en un contexto académico o en un contexto no académico.
- El usuario y el contexto determinan el formato, la naturaleza y la extensión del reporte de investigación.
- Los elementos más comunes de un reporte de investigación presentado en un reporte académico son: portada, índice, resumen, cuerpo del documento (introducción, marco teórico, método, resultados), conclusiones, referencias o bibliografía y apéndices.
- Los elementos más comunes en un contexto no académico son: portada, índice, resumen ejecutivo, método, resultados, conclusiones y apéndices.
- Para presentar el reporte de investigación se pueden utilizar diversos apoyos o recursos.





CONCEPTOS BÁSICOS

Contexto académico
Contexto no académico
Cuerpo del documento

Reporte de investigación
Usuario/receptor

EJERCICIOS

1. Elabore el índice de una tesis.
2. Localice un artículo de una revista científica mencionada en el apéndice 1 del CD anexo y analice los elementos del artículo.
3. Piense en cuál sería el índice del reporte de la investigación que ha concebido a lo largo de los ejercicios del libro y desarróllelo.
4. Elabore una presentación de su tesis o de cualquier investigación realizada por usted u otra persona en un programa para tal efecto disponible en su institución (por ejemplo: Power Point o Flash).



LA TELEVISIÓN Y EL NIÑO

Índice del reporte de investigación

	Página
1. INTRODUCCIÓN	1
1.1 Problema a investigar	2
1.2 Importancia del estudio	5
1.3 Definición de términos	7
1.4 Problemas y limitaciones	10
1.5 Hipótesis	12
2. MARCO TEÓRICO	13
2.1 El enfoque de usos y graficaciones en la comunicación colectiva.....	14
2.2 El uso que los niños hacen de la televisión.....	22
2.3 Contenidos televisivos preferidos por los niños	26
2.4 Las funciones y graficaciones de la televisión para el niño	29
2.5 Elementos que mediatizan las condiciones a las que se exponen los niños al ver televisión.....	37
3. MÉTODO.....	43
3.1 Muestra.....	44
3.2 Contexto y diseño.....	47
3.3 Instrumento de medición	49
3.4 Procedimientos.....	51
3.4.1 Selección de la muestra	51
3.4.2 Recolección de los datos	54
3.4.3 Análisis de los datos	55

4. RESULTADOS	56
4.1 Características demográficas de la muestra	57
4.2 Fuentes alternativas de entretenimiento	60
4.3 Tiempo que dedican los niños a ver la televisión.....	63
4.4 Programas preferidos por los niños	65
4.5 Personajes favoritos	69
4.6 Funciones y gratificaciones de la televisión para los niños de la muestra	73
4.7 Control de los padres	77
5. CONCLUSIONES Y RECOMENDACIONES.....	79
5.1 Resultados fundamentales.....	80
5.2 Conclusiones y discusión.....	82
5.2.1 Implicaciones para los padres.....	84
5.2.2 Implicaciones para los educadores.....	88
5.2.3 Implicaciones para los productores.....	93
5.3 Recomendaciones.....	97
5.4 El futuro de la televisión infantil.....	101
BIBLIOGRAFÍA	105
APÉNDICE A: Carta a los directores de la escuela.....	111
APÉNDICE B: Cuestionario aplicado.....	112

EL CLIMA ORGANIZACIONAL (REPORTE PRIMERAS DOS MUESTRAS)

CONTENIDO	PÁGINA
ÍNDICE I	I
ÍNDICE II	II
RESUMEN	III
SUMMARY	IV
INTRODUCCIÓN	1
PRIMER CAPÍTULO: MARCO TEÓRICO	6
1. Primeros antecedentes	7
2. Definiciones, características y enfoques del clima organizacional	15
3. Dimensiones del clima organizacional	30
4. Modelos del clima organizacional	47
5. Conclusiones al marco teórico	59
SEGUNDO CAPÍTULO: MÉTODO	62
1. Sujetos (muestras)	63
2. Herramientas	63
3. Procedimiento	74
TERCER CAPÍTULO: RESULTADOS	75
1. Coreografía del análisis	76
2. Primera muestra (Laboratorio)	76

2.1. Confiabilidad de toda la escala	76
2.2. Estructura interna de la escala (análisis de factores)	76
2.3. Confiabilidad de la escala tomando en cuenta el análisis de factores	80
2.4. Confiabilidad de las dimensiones del clima organizacional	80
2.5. Correlaciones entre las dimensiones del clima organizacional	81
2.6. Estadística descriptiva al nivel de toda la escala	84
2.7. Estadística descriptiva de las dimensiones del clima organizacional	85
3. Segunda muestra (Institución educativa)	91
3.1. Confiabilidad de toda la escala	91
3.2. Estructura interna de la escala (análisis de factores)	91
3.3. Confiabilidad de la escala tomando en cuenta el análisis de factores	94
3.4. Confiabilidad de las dimensiones del clima organizacional	94
3.5. Correlaciones entre las dimensiones del clima organizacional	95
3.6. Estadística descriptiva al nivel de toda la escala	96
3.7. Estadística descriptiva de las dimensiones del clima organizacional	98
4. Validación de criterio concurrente	102
4.1. Análisis de factores con las escalas totales de las muestras	102
4.2. Regresiones	102
DISCUSIÓN, CONCLUSIONES Y RECOMENDACIONES	105
BIBLIOGRAFÍA	114
APÉNDICES	129
Apéndice número 1: Cuestionario	130
Apéndice número 2: Análisis de factores primera muestra (Laboratorio)	138
Apéndice número 3: Análisis de factores segunda muestra (Institución educativa)	145
Apéndice número 4: Resultados del análisis de regresión.	150

EL ABUSO SEXUAL INFANTIL

Se incluye en el capítulo 9 del CD anexo: "Elaboración de propuestas cuantitativas, cualitativas y mixtas".

A investigar se aprende investigando; por lo tanto, es necesario desmitificar la complejidad de la tarea y sentir pasión por ella. En este sentido, la experiencia en la investigación enriquece ampliamente la labor del docente.

Una investigación será mucho más factible, si el planteamiento del problema se

realiza de manera adecuada; también es importante que el tema sea de actualidad y pertinente, y que esté enfocado a la solución de problemas concretos.

La realidad es cuantitativa-cualitativa; por ello, es necesario combinar ambos enfoques, siempre y cuando no sean incompatibles con el método empleado.

Respecto de la investigación que se realiza en Colombia, de acuerdo con Colciencias, organismo estatal para las ciencias y la tecnología, la Universidad de Antioquia ocupa el primer lugar en todo el país.

Duván Salavarría T.
Profesor-investigador
Facultad de Administración
Universidad de Antioquia
Medellín, Colombia

Una investigación exitosa, es decir, que contribuya de manera trascendente a la generación de conocimiento, depende en gran medida de que el planteamiento del problema se realice adecuadamente.

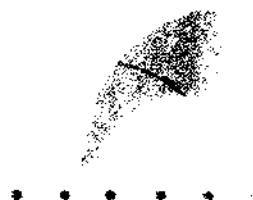
Otro aspecto de consideración es que la investigación puede abarcar tanto el enfoque cualitativo como el cuantitativo, y llegar a complementarse, además de que es posible mezclarlos cuando se utilizan diversos tipos de instrumentos de medición, como registros observacionales, cuestionarios, tests, estudios de caso, etc. En cuanto a paquetes de análisis, en investigación cualitativa actualmente utilizo el SPSS.

Para los estudiantes, la importancia de la investigación radica en que es un medio que brinda la oportunidad de resolver problemas reales, como los que encontrarán en su vida profesional; por ello, la importancia de que elijan un tema de su interés, que además sea original, viable, preciso y de extensión acotada.

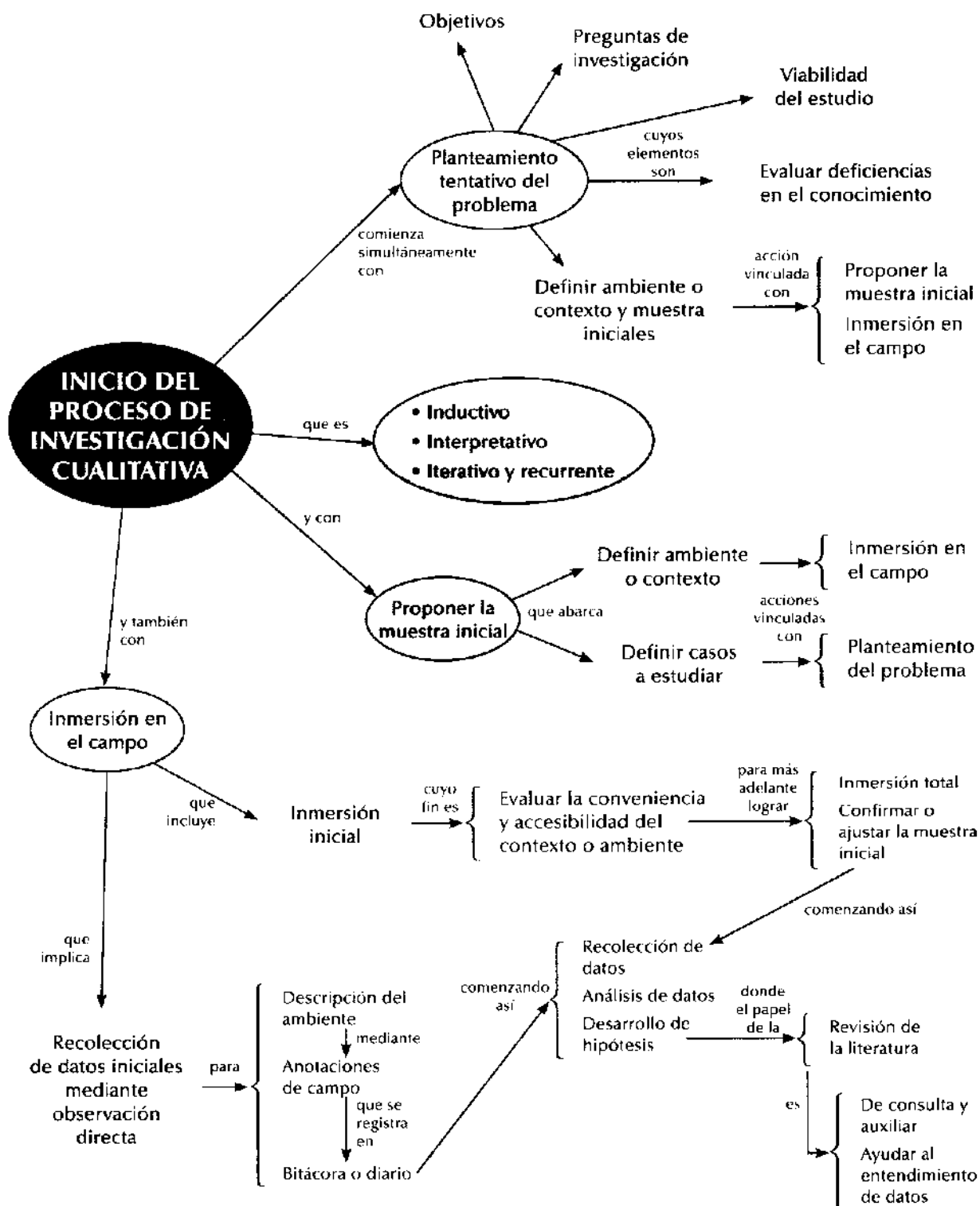
Asimismo, deben tomar en cuenta los parámetros que caracterizan a una buena investigación, y también plantear de forma adecuada el problema. Es necesario definir objetivos precisos; efectuar una intensa revisión bibliográfica; seleccionar el diseño de investigación adecuado; realizar un buen análisis estadístico, el cual representa una herramienta que permite hacer inferencias significativas respecto de los resultados obtenidos; y, por último, llegar a conclusiones objetivas.

Esteban Jaime Camacho Ruiz
Catedrático
Departamentos de Psicología y
Pedagogía
Universidad Hispanoamericana
Estado de México,
México

Tercera parte

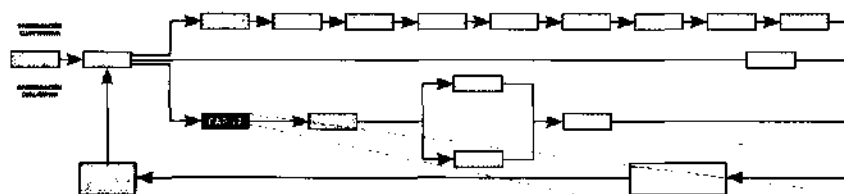


El proceso de la investigación cualitativa



Capítulo 12

El inicio del proceso cualitativo: planteamiento del problema, revisión de la literatura, surgimiento de las hipótesis e inmersión en el campo



OBJETIVOS DE APRENDIZAJE

Al terminar este capítulo, el alumno será capaz de:

- Formular planteamientos para investigar de manera inductiva.
- Visualizar los aspectos que debe tomar en cuenta para iniciar un estudio cualitativo.
- Comprender cómo se inicia una investigación cualitativa.
- Conocer el papel que juegan la revisión de la literatura y las hipótesis en el proceso de investigación cualitativa.

Síntesis

En el presente capítulo se comentará la manera en que la idea se desarrolla y se transforma en el planteamiento del problema de investigación (cualitativo). Es decir, el capítulo trata sobre cómo plantear un problema de investigación, pero ahora desde la óptica cualitativa. Seis elementos resultan fundamentales para plantear un problema cualitativo: objetivos de investigación, preguntas de investigación, justificación de la investigación, viabilidad de ésta, evaluación de las deficiencias en el conocimiento del problema y definición inicial del ambiente o contexto. Sin embargo, los objetivos y las preguntas son más generales y su delimitación es menos precisa. En el capítulo se analizan estos elementos bajo el enfoque cualitativo. Asimismo, se explica el papel que juegan la literatura y las hipótesis en el proceso inductivo; del mismo modo, cómo se inicia, en la práctica, un estudio cualitativo, mediante el ingreso al contexto, ambiente o campo.

Por otro lado, se insiste en que el proceso cualitativo no es lineal, sino iterativo o recurrente, las supuestas etapas en realidad son acciones para adentrarnos más en el problema de investigación y la tarea de recolectar y analizar datos es permanente.

PROCESO DE INVESTIGACIÓN CUALITATIVA

Paso 2 Planteamiento del problema

- Establecer objetivos y preguntas de investigación iniciales, justificación y viabilidad.
- Definir tentativamente el papel que desempeñará la literatura.
- Elegir el ambiente o contexto donde se comenzará a estudiar el problema de investigación.
- Entrar en el ambiente o contexto.

... ¿Qué es plantear el problema de investigación cualitativa?

Una vez concebida la idea del estudio, el investigador debe familiarizarse con el tema en cuestión. Aunque el enfoque cualitativo es inductivo, necesitamos conocer con mayor profundidad el “terreno que estamos pisando”. Imaginemos que estamos interesados en realizar una investigación sobre una cultura indígena, sus valores, ritos y costumbres. En este caso debemos saber al menos dónde radica tal cultura, su antigüedad, sus características esenciales (actividades económicas, religión, nivel tecnológico, total aproximado de su población, etc.) y qué tan hostil es con los extraños. De igual forma, si pretendemos estudiar las experiencias de vida de diferentes personas en torno a un hecho (una guerra, por ejemplo), requerimos tener datos sobre cuándo ocurrió esa guerra, quiénes la protagonizaron y qué consecuencias tuvo. Asimismo, si vamos a estudiar la depresión posparto en ciertas mujeres, es necesario que tengamos conocimiento de qué la distingue de otros tipos de depresión y cómo se manifiesta.

Ya que nos hemos adentrado en el tema, podemos plantear nuestro problema de estudio. El planteamiento cualitativo suele incluir los objetivos, las preguntas de investigación, la justificación y la viabilidad, además de una exploración de las deficiencias en el conocimiento del problema y la definición inicial del ambiente o contexto.

Los objetivos y preguntas poseen una naturaleza distinta en la investigación cualitativa, ya que en ésta no se busca precisar ni acotar el problema o fenómeno de estudio al comenzar el proceso. Tales elementos del planteamiento son más bien enunciativos y constituyen un punto de partida en la investigación.

La justificación es importante, al igual que en los planteamientos cuantitativos, particularmente cuando el estudio necesita de la aprobación de otras personas. Finalmente, requieren una razón para llevarse a cabo; y una vez más aparecen los criterios ya comentados en el capítulo 3 del libro: conveniencia, relevancia social, implicaciones prácticas, valor teórico y utilidad metodológica. Asimismo, en la justificación se pueden incluir datos cuantitativos para dimensionar el problema de estudio, aunque nuestro abordaje sea cualitativo. Si la investigación es sobre las consecuencias del abuso sexual infantil, el planteamiento puede enriquecerse con datos y testimonios (por ejemplo, estadísticas sobre el número de abusos reportados, sus consecuencias y daños).

La viabilidad es un elemento que también se valora y se ubica en cuanto a tiempo, recursos y habilidades. Es necesario que nos cuestionemos: ¿es posible llevar a cabo el estudio?, ¿poseemos los recursos para hacerlo?

En relación con las deficiencias en el conocimiento del problema, constituyen un elemento que se considera principalmente si se decide revisar la literatura relacionada con el tema de la investigación. Sobre este tópico comentaremos más adelante en este capítulo. Puede ocurrir que se pretenda plantear un problema cualitativo cuando hay poca investigación inductiva respecto a un tema; esto es, que se han realizado diversos estudios cuantitativos, pero no se indagado el fenómeno de manera cualitativa o se requiere profundizar en ciertos resultados de tales estudios.

Si los planteamientos cuantitativos pueden modificarse, los cualitativos con más razón. Su flexibilidad es aún mayor. Éstos últimos, en lugar de establecer variables “exactas”, identifican conceptos esenciales para iniciar la investigación.

Williams, Unrau y Grinnell (2005) establecen una excelente metáfora de lo que representa un planteamiento cualitativo: es como entrar a un laberinto, sabemos dónde comenzamos, pero no dónde habremos de terminar. Entramos con convicción, pero sin un “mapa” preciso.

Una comparación entre planteamientos cuantitativos y cualitativos puede ayudar a reforzar los puntos anteriores (ver tabla 12.1).

Un ejemplo de un planteamiento cualitativo podría ser el que se comenta a continuación. Supongamos que nos interesa efectuar una investigación sobre las emociones que pueden experimentar los pacientes jóvenes que serán intervenidos en una operación de tumor cerebral. El planteamiento podría ser:

Objetivos:

1. Conocer las emociones que experimentan pacientes jóvenes que serán intervenidos en una operación de tumor cerebral.
2. Profundizar en las vivencias de tales pacientes y su significado.
3. Comprender los mecanismos que el paciente utiliza para confrontar las emociones negativas profundas que surgen en la etapa preoperatoria.

Preguntas de investigación:

1. ¿Qué emociones experimentan los pacientes jóvenes que serán intervenidos en una operación de tumor cerebral?
2. ¿Cuáles son sus vivencias antes de ser intervenidos quirúrgicamente?
3. ¿Qué mecanismos utilizan para confrontar las emociones negativas que surgen en la etapa previa a la operación?

Tabla 12.1 Comparación entre planteamientos cuantitativos y cualitativos

Planteamientos cuantitativos	Planteamientos cualitativos
Precisos y acotados o delimitados	Abiertos
Enfocados en variables lo más exactas y concretas que sea posible	Expansivos, que paulatinamente se van enfocando en conceptos relevantes de acuerdo con la evolución del estudio
Direccionados	No direccionados en su inicio
Fundamentados en la revisión de la literatura	Fundamentados en la experiencia e intuición
Se aplican a un gran número de casos	Se aplican a un menor número de casos
El entendimiento del fenómeno se guía a través de ciertas dimensiones consideradas como significativas por estudios previos	El entendimiento del fenómeno es en todas sus dimensiones, internas y externas, pasadas y presentes
Se orientan a probar teorías, hipótesis y/o explicaciones, así como a evaluar efectos de unas variables sobre otras (los correlacionales y explicativos)	Se orientan a aprender de experiencias y puntos de vista de los individuos, valorar procesos y generar teorías fundamentadas en las perspectivas de los participantes

Ahora bien, para responder a las preguntas es necesario elegir un contexto o ambiente donde se lleve a cabo el estudio, pues aunque los planteamientos cualitativos son más generales, deben situarnos en tiempo y lugar (Creswell, 2005). En el planteamiento mencionado: ¿qué sitio es el adecuado? La lógica nos indica que uno o varios hospitales son el contexto apropiado para nuestra investigación. Por otro lado, nos interesa analizar a jóvenes; pero, ¿cuáles jóvenes? (rusos, italianos, españoles, mexicanos, romanos, vascos, regiomontanos). Es claro que entonces debemos definir el intervalo que para nosotros abarca el concepto de jóvenes (supongamos que el intervalo incluye a personas de 13 a 17 años) y la provincia o ciudad (por ejemplo: Salta, en Argentina). Además, nos interesan las experiencias de jóvenes que serán intervenidos en una operación de tumor cerebral (hemos desechado a quienes se someten a intervenciones quirúrgicas menores u otro tipo de operación).

Lo primero es obtener información sobre cuáles hospitales de la ciudad realizan operaciones de esta naturaleza con regularidad. Pudiera ocurrir que al acceder a los registros, éstos nos indican que en todos los hospitales de la ciudad efectúan esta operación, pero sólo una vez a la semana. Esto implicaría que realizar el estudio puede llevar una considerable cantidad de tiempo. Podemos decidir que esto no es relevante en nuestro caso y proseguir. O bien, que debemos ampliar nuestro rango de edades o de tipo de operaciones. Otro panorama sería que, desafortunadamente para los jóvenes, esta operación ocurre con mayor frecuencia.

Al plantear el problema, es importante tener en mente que la investigación cualitativa: a) es conducida primordialmente en los ambientes naturales de los participantes (en este caso, hospitales, desde el cuarto del paciente y la zona preoperatoria hasta el restaurante del hospital y los corredores o pasillos); b) las variables no son controladas ni manipuladas (incluso no definimos variables sino conceptos generales como “emociones”, “vivencias” y “mecanismos de confrontación”); c) los significados serán extraídos de los participantes, y d) los datos no se reducirán a valores numéricos (Rothery, Tutty y Grinnell, 1996).

Una vez hecho el planteamiento inicial empezáramos a contactar a los participantes potenciales y a recolectar datos, probablemente el método que utilicemos para esta labor sea la entrevista. Efectuada la primera entrevista podríamos comenzar a generar datos y tal vez nos percatemos de que los jóvenes antes de ingresar al quirófano experimentan un elevado estrés. En otras entrevistas podríamos seguir detectando ese estrés y enfocarnos en él. Los datos nos movilizan en diferentes direcciones y es cuando vamos respondiendo al problema original y modificándolo.

En cuanto a si debemos o no revisar literatura al plantear el problema de estudio cualitativo, tres posiciones se pueden identificar de los autores de obras sobre metodología de la investigación:

1. La posición que denominamos “radical”, que indica no revisar investigaciones anteriores y mucho menos tomarlas en cuenta.
2. La que etiquetamos como posición “intermedia”, la cual sugiere revisar estudios previos únicamente para: a) identificar conceptos que nos puedan ser útiles para elaborar nuestro planteamiento y b) ayudarnos a visualizar cómo han sido descritos y definidos tales conceptos por otros investigadores.
3. La posición que podría denominarse como “integradora”, considera la conveniencia de revisar los trabajos previos con el fin de que contribuyan en la construcción del planteamiento e incluso para elaborar una perspectiva teórica, pero finalmente desprenderse de ésta al inducir y sólo mantenerla como una referencia.

Ciertamente, la posición radical es extrema y puede llevarnos a cometer errores en el proceso investigativo, aun cuando creamos conocer con profundidad el fenómeno que estamos indagando. La humildad siempre será una cualidad deseable en la investigación científica.

Las otras dos posiciones, más razonables desde nuestro punto de vista, implican valorar la investigación hecha por otros y ser precavidos en nuestro camino por encontrar respuestas a las interrogantes planteadas, lo cual significa que la literatura puede sernos útil para distintos propósitos:

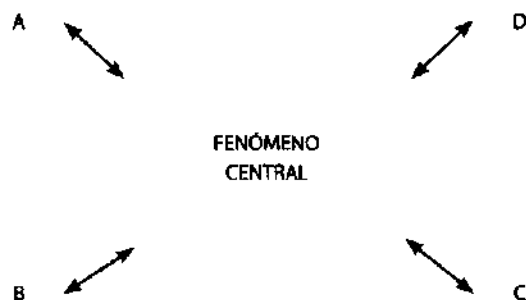
1. Detectar conceptos clave que no habíamos pensado.
2. Nutrirnos de ideas en cuanto a métodos de recolección de datos y análisis, respecto de cómo les han servido a otros.
3. Tener en mente los errores que otros han cometido anteriormente.
4. Conocer diferentes maneras de pensar y abordar el planteamiento.

Dejar a un lado “el pasado” es algo ingenuo e irreal, pues como mencionan Williams, Unrau y Grinnell (2005), siempre comenzamos una investigación con ciertas experiencias, ideas y opiniones sobre el problema a estudiar, lo cual es resultado de nuestra propia historia de vida. Desde luego, tratamos de hacer a un lado –en la medida de lo posible– nuestras opiniones sobre cómo se relacionan los conceptos, de igual forma nos mantenemos abiertos a nuevos conceptos y a las relaciones que emerjan entre éstos.

Volviendo a la metáfora del laberinto: la revisión de la literatura nos puede al menos decir si el laberinto se ha percibido como largo o corto, como intrincado o menos complejo; si tiene esquinas circulares o perpendiculares; si tiene o no varios niveles. Desde luego, para movernos en el laberinto, nos basamos en nuestras experiencias en el momento que transitamos por él. Como, finalmente, nunca hemos estado en el laberinto, debemos caminar con precaución observando todo, utilizando cada uno de nuestros sentidos y confiando también en la propia intuición, pero también debemos estar preparados para cualquier contratiempo (podríamos entrar en zonas sin salida y tener que regresarnos, encontrar cruces no imaginados y hasta descubrir pasadizos secretos). Obviamente, tenemos que estar dispuestos a renunciar a lo que creíamos saber de laberintos y aceptar nuevas experiencias.

En resumen, la *revisión de la literatura* puede servirnos en el planteamiento del problema cualitativo inicial; pero a diferencia de los planteamientos cuantitativos –no tanto los exploratorios, como los descriptivos, correlaciones o explicativos–, nuestro fundamento no se circunscribe o limita a dicha revisión, incluso, su papel es más bien de apoyo. La *investigación cualitativa* se basa, ante todo, en el proceso mismo de recolección y análisis. Recordemos que es interpretativa, ya que el investigador hace su propia descripción y valoración de los datos. El planteamiento se va enfocando en ciertos temas de acuerdo con la información recabada.

Una manera que sugerimos para comenzar a plantear el *problema de investigación*, se muestra en la figura 12.1. El procedimiento es muy sencillo: primero, definimos el concepto central de nuestro estudio y los conceptos que consideramos se vinculan con él, de acuerdo con nuestra experiencia y la revisión de la literatura. Posteriormente, volvemos a revisar el esquema a lo largo de la indagación y lo vamos consolidando, precisando o modificando conforme recogemos y evaluamos los datos. Veamos un caso ilustrativo de ello.



A, B, C y D representan los conceptos vinculados.
Por ejemplo:

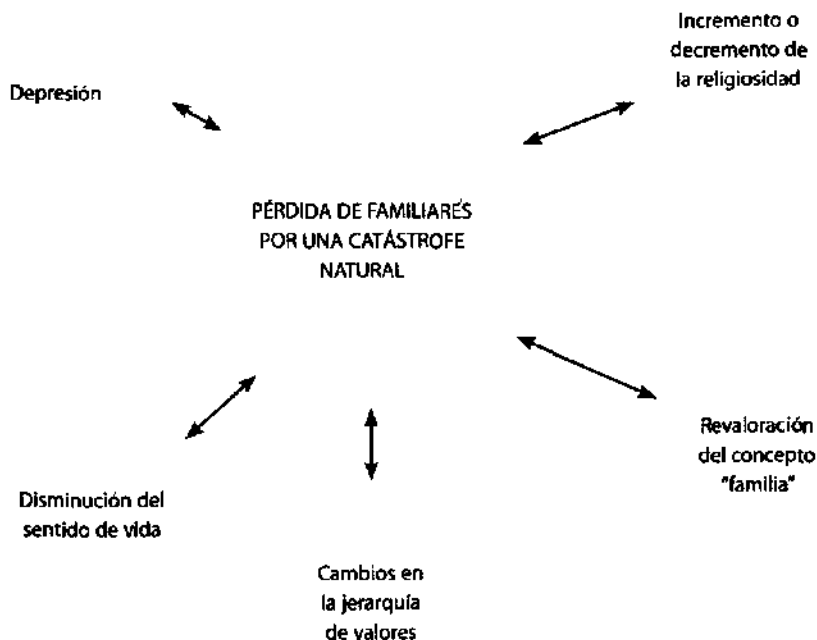


Figura 12.1 Sugerencia para la visualización gráfica de un planteamiento cualitativo.

El interés del estudio podría ser general, por ejemplo, entender profundamente la experiencia humana que significa perder a un familiar a consecuencia de un desastre natural (un terremoto, una inundación, un *tsunami*, etc.). Este es el concepto central. Entonces el planteamiento inicial sería tan genérico como:

Objetivo: Entender el significado de la experiencia humana resultante de la pérdida de un familiar a consecuencia de un desastre natural.

Pregunta de investigación: ¿Cuál es el significado que tiene para un ser humano la pérdida de un familiar a consecuencia de un desastre natural?

El porqué estamos interesados en una investigación así, complementaría el planteamiento junto con la viabilidad del estudio.

Justificación (en términos resumidos): Al entender el significado de tales experiencias y la realidad personal de los individuos que las viven, podemos obtener un conocimiento más profundo de la naturaleza humana en casos de desastre y planear mejores esquemas de apoyo psicológico para sus víctimas. Tal conocimiento nos permite, al menos, una mayor empatía con los seres humanos que sufren la pérdida de un familiar a consecuencia de un fenómeno natural.

Viabilidad: Hace dos días ocurrió un terremoto con consecuencias fatales y puede efectuarse la investigación. Se cuenta con los recursos y conocimientos para ello.

O bien, el planteamiento podría enfocarse en el concepto central y otros conceptos relacionados, extraídos de nuestras reflexiones, experiencias y la revisión de la literatura; como en el ejemplo de la figura 12.1: depresión, disminución en el sentido de vida, cambios en la jerarquía de valores (reposicionamiento de valores humanos colectivos como la solidaridad, la convivencia, etc.), revaloración del concepto "familia" e incremento o decremento en la religiosidad (mayor apego a las creencias religiosas o al contrario, su pérdida). Así, el planteamiento podría ser:

Objetivo: Entender el significado de la experiencia humana resultante de la pérdida de un familiar a consecuencia de un desastre natural y su relación con la depresión, la disminución en el sentido de vida, los cambios en la jerarquía de valores, la revaloración del concepto "familia" y el incremento o decremento en la religiosidad.

Pregunta de investigación: ¿Cuál es el significado que tiene para un ser humano la pérdida de un familiar como resultado de un desastre natural y cómo se vincula con la depresión, la disminución en el sentido de vida, los cambios en la jerarquía de valores, la revaloración del concepto "familia" y el incremento o decremento en la religiosidad?

Incluso podría enfocarse únicamente en la depresión que origina tal categoría de tragedias. Es decir, el planteamiento puede ser más o menos general, y debe ubicarse en un contexto, en este caso un desastre natural concreto (como por ejemplo, el maremoto o *tsunami* en el océano Índico que en diciembre de 2004 devastó parte de Asia, o el huracán Katrina que destruyó ciudades y poblados en el sureste de los Estados Unidos de Norteamérica en agosto de 2005). Un ejemplo de una investigación cualitativa posterior a un desastre natural (con niños, aunque no necesariamente habían perdido a un familiar) se realizó en la Escuela de Psicología de la Universidad de Colima en 2003 (Montes, Otero, Castillo y Álvarez, 2003), después de un severo terremoto de 7.8 grados en la escala de Richter que sacudió la zona donde se ubica dicha institución. Primero, se documentaron experiencias emocionales de niños ante el temblor y se les proporcionó intervención psicológica; después se elaboró un programa para difundir una cultura de prevención de desastres, dirigido a los infantes de escuelas primarias de la Ciudad de Colima, México.

Los resultados de este tipo de estudios no intentan generalizarse a poblaciones más amplias, sino que se dirigen a la comprensión de vivencias en un entorno específico, cuyos datos emergentes aportan al entendimiento del fenómeno.

Con un esquema como el de la figura 12.1 no se pretende estandarizar la generación de planteamientos cualitativos, es simplemente una sugerencia respecto de las múltiples formas de visualizarlos. Creswell (2005) nos recomienda otra forma gráfica para plantear problemas cualitativos (ver figura 12.2).

Los planteamientos cualitativos son una especie de plan de exploración (entendimiento emergente) y resultan apropiados cuando el investigador se interesa por el significado de las experiencias y valores humanos, el punto de vista interno e individual de las personas y el ambiente natural en que ocurre el fenómeno estudiado, así como cuando buscamos una perspectiva cercana de los participantes. Patton (2002) identifica las siguientes áreas y necesidades como adecuadas para planteamientos cualitativos referentes a procesos (por ejemplo, en torno a un programa educativo o uno de cambio organizacional):

1. El centro de la investigación está conformado por las experiencias de los participantes en torno al proceso, particularmente si subraya resultados individualizados.
2. Es necesaria información detallada y profunda acerca del proceso.
3. Se busca conocer la diversidad de idiosincrasias y cualidades únicas de los participantes inmersos en el proceso.

Mertens (2005), además de Coleman y Unrau (2005) consideran que la investigación cualitativa es particularmente útil cuando el fenómeno de interés es muy difícil de medir o no se ha medido anteriormente (deficiencias en el conocimiento del problema). Tal fue el caso de un estudio, donde Donna Mertens y otros colaboradores pretendieron evaluar el impacto de la sensibilización —vía entrenamiento— sobre las actitudes de maestros y administradores egipcios hacia personas con capacidades distintas. Al no encontrar instrumentos estandarizados en la cultura egipcia, prefirieron recolectar datos mediante técnicas cualitativas (observaciones y entrevistas que además, documentaron el lenguaje empleado para describir a dichas personas). Otro caso lo sería un estudio para profundizar en el miedo que experimentan ciertas mujeres al ser agredidas físicamente por sus esposos. En situaciones como éstas, la cuantificación incluso podría resultar trivial. Sería más adecuado adentrarse en el significado profundo de la experiencia de las mujeres.

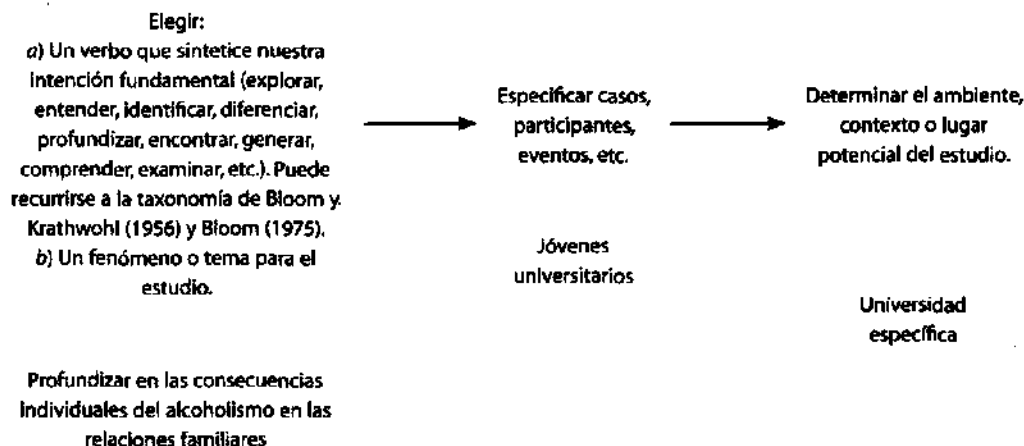


Figura 12.2 Otro modelo para el planteamiento de problemas cualitativos.

En resumen, ya sea que se trate de planteamientos cualitativos con mayor o menor enfoque, el punto de partida será, prácticamente en la mayoría de los casos, la presencia del investigador en el contexto, donde comienza su inducción. El grado en que utilicemos a la literatura para elaborar el planteamiento es decisión del investigador en cada caso.

¿Qué papel desempeña el marco teórico en la investigación cualitativa?

Ya se mencionaron las posiciones más relevantes respecto del papel que desempeña la literatura en el planteamiento del problema, las cuales pueden extenderse a todo el proceso de la investigación cualitativa. Lo importante, a nuestro juicio, es que en cualquier caso la literatura no debe “empañar” la tarea inductiva.

En los estudios cualitativos, más que integrarse un marco teórico formal, como lo es en el caso de la investigación cuantitativa, se revisa la literatura con los propósitos comentados previamente y su papel varía en el inicio y la parte final del estudio. Creswell (2005, p. 80) compara el papel que posee la literatura en la investigación cuantitativa y la cualitativa de la forma en que se muestra en la tabla 12.2.

En diversos estudios cualitativos, la literatura se revisa y se mantiene “conceptualmente” distante al comenzar a recolectar datos y posteriormente se va consultando de acuerdo con la evolución de la investigación, en este sentido, su función esencial –tal como se comentó– es de consulta. En los estudios cuantitativos, el papel de la teoría es establecer relaciones entre variables o constructos que describen y explican un fenómeno. En la indagación cualitativa su papel auxiliar implica proveer de ideas no contempladas, pero que, desde luego, se ajustan al contexto y desarrollo del estudio; que al final, le permite al investigador “ir más allá del evento local”, al relacionarlo con eventos similares. La literatura colabora a mejorar el entendimiento de los datos

Marco teórico en estudios cualitativos No prefigura con exactitud el planteamiento del problema.

Tabla 12.2 Diferencias en la extensión y uso de la literatura en la investigación cuantitativa y cualitativa

Diferencia	Investigación cuantitativa	Investigación cualitativa
Cantidad de literatura citada al comienzo del estudio.	Sustancial.	Mínima. Normalmente no se discute extensivamente la literatura para facilitar que los datos o información emerjan de los participantes, sin limitarnos a la visión de otros estudios.
Utilización o funciones de la literatura al inicio del estudio.	Proveer una dirección racional al estudio (por ejemplo, afinar el planteamiento e hipótesis).	Justificar y documentar la necesidad de realizar el estudio.
Utilización de la literatura al final del estudio.	Confirmar o no las predicciones previas emanadas de la literatura.	Apoyar o modificar los resultados existentes documentados en la literatura.

recolectados y analizados, pero siempre el investigador se orienta fundamentalmente por los resultados que emergen del trabajo en el contexto o ambiente particular. Es bastante común revisar la literatura de vez en cuando para profundizar las interpretaciones (Mertens, 2005). Y cabría agregar que entre la literatura y los datos emergentes, éstos tienen “la última palabra”. Un ejemplo ilustrativo del típico uso de la literatura en una investigación cualitativa podría ser el que se muestra en la figura 12.3.

De esta forma, sin romper con nuestro esquema inductivo, podemos apoyarnos en datos estadísticos y otros estudios para dimensionar el problema en Valledupar; utilizar las investigaciones previas para mejorar la recolección y análisis de los datos, y al final, tener un punto con el cual contrastar los resultados, identificar los estudios que han llegado a conclusiones similares, pero no en un sentido predictivo como en la investigación cuantitativa, sino en términos de ideas y prácticas (Creswell, 2005). Es decir, en cualquier etapa podemos consultar la literatura y constantemente nos sumergimos en el campo.

Finalmente, la decisión de qué tanto se involucra a los estudios antecedentes como un punto de referencia es del investigador.

Preguntas de investigación:

¿Cómo pueden describirse las experiencias de ciertas mujeres de Valledupar que son agredidas físicamente por sus esposos? ¿Qué provoca estas agresiones? ¿Por qué dichas mujeres mantienen la relación marital?

Profundizar en las consecuencias individuales del alcoholismo en las relaciones familiares.

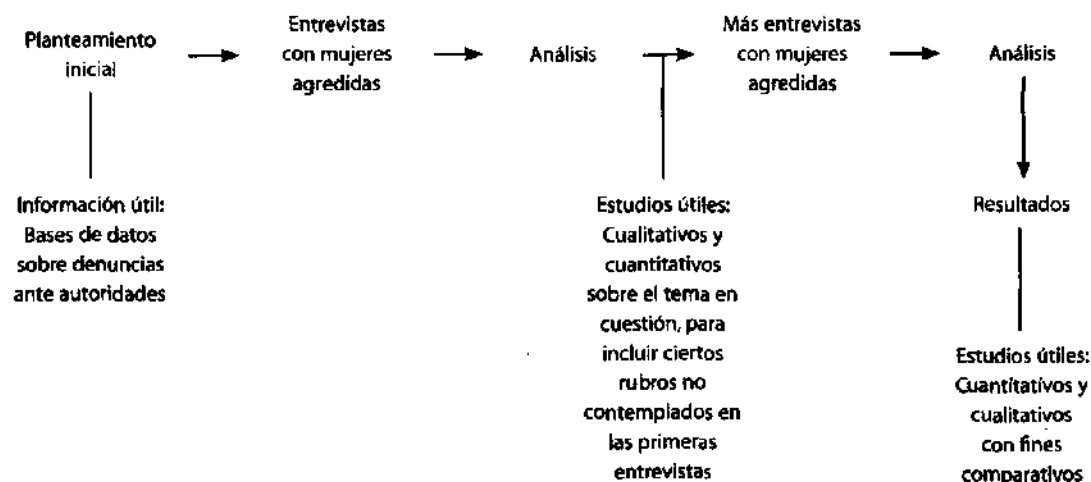


Figura 12.3 Empleo de la literatura en un problema específico de investigación cualitativa.

¿Qué papel desempeñan las hipótesis en el proceso de investigación cualitativa?

En los estudios cualitativos, las hipótesis adquieren un papel distinto al que tienen en la investigación cuantitativa. En primer término, en raras ocasiones se establecen antes de ingresar en el ambiente o contexto y comenzar la recolección de los datos (Williams, Unrau y Grinnell, 2005). Más bien, durante el proceso, el investigador va generando hipótesis de trabajo que se afinan paulatinamente conforme se recaban más datos, o las hipótesis son uno de los resultados del estudio. Las hipótesis se modifican sobre la base de los razonamientos del investigador y, desde luego, no se prueban estadísticamente.

Las **hipótesis de trabajo** cualitativas son pues, emergentes, flexibles y contextuales, se adaptan a los datos y avatares del curso de la investigación.

Por ejemplo, en un estudio sobre las oportunidades de empleo para las personas con capacidades diferentes en un municipio de medio millón de habitantes (Amate y Morales, 2005), se comenzó con la idea de que tales oportunidades eran desfavorables para dichas personas. Sin embargo, al comenzar a observar lo que sucedía en algunas empresas y entrevistar a directores de las áreas de recursos humanos o administración de personal, así como a obreros, se pudo determinar que la idea inicial era incorrecta: que las oportunidades eran iguales para individuos con capacidades regulares que para aquellos con capacidades distintas. Esta hipótesis de trabajo fue variando conforme se recogieron más datos, hasta que se concluyó que: "Las empresas transnacionales o con presencia en todo el país son organizaciones que ofrecen oportunidades similares tanto a las personas con capacidades regulares, como a los individuos con capacidades diferentes; porque poseen recursos para ofrecerles a estas últimas entrenamiento en cualquier actividad laboral. Pero las empresas locales carecen de tales recursos y no ofrecen oportunidades iguales, la cuestión no tenía que ver con prejuicio o discriminación, sino con posibilidades económicas (querían, pero no podían)".

Una vez hecho el planteamiento inicial y definido el papel de la literatura, ¿qué sigue?

El ingreso en el ambiente (campo)

Una vez que hemos elegido un ambiente, contexto o lugar apropiado, comenzamos la tarea de responder a las preguntas de investigación. El ambiente puede ser tan variado como el planteamiento del problema (un hospital, una o varias empresas, una zona selvática –si estudiamos el comportamiento de una especie animal–, una comunidad indígena, una universidad, una plaza pública, un consultorio, una casa donde sesiona un grupo, etc.). Y el contexto implica una definición geográfica, pero es inicial, puesto que puede variar, ampliarse o reducirse. Imaginemos que queremos estudiar los valores de ciertos estudiantes universitarios mediante la observación de conductas que los reflejen o representen. El sitio inicial podría ser el campus de una institución, pero después tendríamos que cambiar los lugares de observación (antros, bares y restaurantes a donde acuden, salas cinematográficas, centros deportivos y de entretenimiento, entre otros). Si

la investigación es sobre pandillas, tendremos que acudir a los puntos donde se reúnen y lo sitios donde actúan. En el caso del estudio de Montes *et al.* (2003), el contexto fue el de las escuelas primarias.

Un tipo de estudios muy socorrido es el denominado “clientes misteriosos” (*mystery shoppers*), donde personas que son supuestos clientes (pero que en realidad son evaluadores calificados) valoran niveles de servicio en la atención (se construyen casos o situaciones específicas para analizar cuestiones como tiempos de espera en el servicio, amabilidad del personal que trata con el cliente, resolución de problemas, manejo de clientes difíciles). En tales investigaciones el ambiente puede constituirse por todos los lugares donde se tiene contacto con el cliente; por ejemplo, en un hotel se abarcaría desde el estacionamiento, el vestíbulo y la recepción, hasta el restaurante y demás espacios como las habitaciones o cuartos, los elevadores, pasillos, etcétera.

La primera tarea es **explorar el contexto** que se seleccionó inicialmente. Lo que significa visitarlo y evaluarlo para cerciorarnos que es el adecuado. Incluso, para considerar nuestra relación con el ambiente por medio de una serie de reflexiones y resolver cualquier situación que pueda entorpecer el estudio (Esterberg, 2002):

1. ¿Me conocen en dicho ambiente?, en caso de que me conozcan los participantes: ¿cómo puedo manejarlo sin afectar a la investigación?
2. Soy muy distinto a los participantes del estudio y mi cotidianidad no tiene que ver con la del ambiente (por ejemplo: pertenecen a un grupo étnico o una clase social muy diferente a la mía), ¿cómo puedo manejarlo? Imaginemos que la investigación es sobre los ritos que un grupo indígena posee para enterrar a sus muertos (para comenzar hablan una lengua distinta a la mía, su cultura es otra y hasta mi físico no es igual). O bien, un estudio sobre los integrantes de una pandilla como los “maras salvatruchas”,¹ con los cuales podemos tener poco o nada en común.
3. ¿Qué significados tiene para mí el contexto?, ¿puedo manejarlos? Por ejemplo, si he tenido experiencias difíciles con los pandilleros (me han asaltado varias veces) y los voy a estudiar o si el ambiente es un hospital donde murió un amigo muy querido.

Asimismo, para estimar tentativamente el tiempo aproximado que nos llevará el estudio y revalorar su viabilidad, porque como menciona Mertens (2005), dos dimensiones resultan esenciales con respecto al ambiente: *conveniencia y accesibilidad*. La primera responde a las siguientes interrogantes: ¿el ambiente definido contiene los casos, personas, eventos, situaciones, historias y/o vivencias que necesitamos para responder a la(s) pregunta(s) de investigación? La segunda tiene que ver con el cuestionamiento: ¿es factible realizar la recolección de los datos? o ¿podemos acceder a los datos que necesitamos? Lograr el acceso al ambiente es una condición para seguir con la investigación e implica obtener permiso de parte de quienes controlan el ingreso (denominados *gatekeepers*).

¹ Grupos de pandillas que se originaron a mediados en los años sesenta en El Salvador y que se han extendido por Centroamérica y Norteamérica. Al inicio se trataba de refugiados que buscaban externalizar su rebeldía. Pero con el tiempo, algunas de las pandillas han evolucionado hasta convertirse en criminales. El nombre de “Mara Salvatrucha” proviene de: “mara”, que se emplea en el Salvador para nombrar a la gente alborotada; “salva”, de salvadoreño y “trucha”, significa listo o espabilado (Bobango, 1981).

Lo anterior significa, sin lugar a dudas, negociar con estas personas (en una empresa pueden ser el director y su gerente de recursos humanos u otros gerentes, en un hospital el director y la junta médica, en una pandilla el líder y su grupo cúpula, en un barrio un presidente de una asociación vecinal o de colonos). Es imprescindible exponerles el estudio, normalmente por medio de una presentación visual y la entrega del proyecto o protocolo (que incluye el planteamiento, por qué fue elegido el ambiente, quiénes serán los participantes, cuánto tiempo aproximadamente pensamos estar en el ambiente o campo, qué se va a hacer con los resultados, dónde se pretende publicarlos, etc.). Asimismo, podemos ofrecerles alguno de los productos o resultados, tales como: un diagnóstico vinculado al planteamiento (de la cultura organizacional, la problemática de las pandillas locales, de una enfermedad), contribuir a la solución de un problema (alcoholismo de jóvenes, capacitación de obreros...), elaborar un plan o manuales (para atender psicológicamente a víctimas de un desastre, mejorar el trato a los pacientes que van a ser intervenidos quirúrgicamente...), etc. A veces la negociación es directa con los participantes (por ejemplo, mujeres agredidas por sus esposos) o una mezcla de éstos y los *gatekeepers*. Es normal que diversas organizaciones, comunidades y personas sean reacias a que otros individuos las observen, ya que el temor a la evaluación es natural.

Asimismo, debemos seleccionar ambientes o lugares alternos, en caso de que el acceso al contexto original nos sea negado o restringido más allá de lo razonable. Desde luego, también debemos visitarlos. Algunas de las recomendaciones para tener un mayor y mejor acceso al ambiente, así como ser aceptados, son las que a continuación se enumeran.

Los *gatekeepers* o controladores de ingreso a un lugar Son individuos que a veces tienen un papel oficial en el contexto y otras veces no, pero de cualquier manera pueden autorizar la entrada al ambiente o al menos facilitarla. También ayudan al investigador a localizar participantes y lo asisten en la identificación de lugares.

1. Desarrollar relaciones

- Ganarnos la confianza de los *gatekeepers* y de los participantes, al ser amables, honestos, sensibles, cooperativos y sinceros (entre más amigos es mejor).
- Apoyarlos en alguna necesidad (gestionar asesoría médica o psicológica, en lo educativo; resolver problemas: hay quien ha arreglado desde un cortocircuito hasta una tubería; dar obsequios, transportar personas: "dar aventón" o "convertirnos en taxi").
- Detectar y cultivar informantes clave (varios para contar con mayor información y diferentes perspectivas (en un hospital, enfermeras, personal de aseo, médicos; en una empresa, trabajadores con bastante antigüedad, la secretaria de tal persona, etc.). En ocasiones podemos aprovechar nuestras redes personales (por ejemplo, en estudios con religiosas, un pariente sacerdote o una monja conocida nos pueden ayudar significativamente; lo mismo con cuerpos policiacos y en hospitales).

2. Elaborar una historia sobre la investigación

- Tener preparado un pequeño guión sobre el estudio (propósito central, tiempo aproximado de estancia en el ambiente, uso de los resultados). Es importante hablar de la investigación, salvo que afecte los resultados, en cuyo caso se recomienda elaborar una versión que sea lo más cercana a la verdad pero no obstrusiva. Por ejemplo, si el estudio es sobre la equidad de género que poseen los profesores en sus relaciones con los alumnos, y les explicamos a todos de

qué se trata la indagación, tal vez su comportamiento deje de ser natural; entonces les podemos comentar que el estudio es sobre actitudes de maestros y alumnos en el aula (simplemente es mejor “disfrazar la situación”) y al concluir el análisis habrá que explicarles la investigación verdadera y sus resultados.

Al final, nunca debemos mentir ni engañar, después de todo: ¿cómo esperar que sean honestos si nosotros no lo somos? Es necesario preparar algunas respuestas para las preguntas que muchas veces suelen inquirirnos los participantes. Por ejemplo: ¿por qué debo cooperar con el estudio?, ¿qué gano yo o los míos con la investigación?, ¿por qué fui elegido para participar en el estudio?, ¿quiénes se benefician con los resultados?

3. *No intentar imitar a los participantes, supuestamente para ganar empatía*

- “No hay algo peor que una persona citadina queriendo actuar como vaquero, ranchero, agricultor o campesino.” Es absurdo y grotesco. Si somos distintos, debemos asumir las diferencias y actuar nuestro papel como investigadores. En todo caso, es preferible agregar al equipo del estudio, a una persona con las mismas características de los participantes y que posea los conocimientos necesarios o prepararlo; o ayudarnos de alguien interno. Esto es bastante común en estudios de etnias. También, puede tenerse un diseño de investigación participativa, en el cual las mismas personas del contexto colaboran en diferentes partes del estudio.

4. *Planear el ingreso al ambiente o contexto (campo)*

- Por lo regular, es mejor entrar de la manera menos disruptiva posible. El ingreso debe resultar natural. Si mantenemos buenas relaciones desde el principio, nos acomodamos a las rutinas de los participantes, establecemos lo que tenemos en común con ellos (que es muy diferente intentar imitarlos), demostramos un genuino interés por la comunidad y ayudamos a la gente, el acceso será “menos ruidoso” y más efectivo. La planeación no es exacta y debemos estar preparados ante cualquier contingencia que suele ocurrir en los estudios cualitativos. A veces el plan de entrada es paulatino (un ingreso por etapas). Por ejemplo, si el estudio es sobre violencia intrafamiliar (padres de familia que agreden a su cónyuge y a sus hijos e hijas), primero debemos plantear una investigación sobre la familia en general (nuestra recolección inicial se orientaría a cuestiones sobre cómo son las familias, cuántos integrantes las componen, qué costumbres tienen, etc.); en una segunda instancia, la indagación se aboca a los problemas familiares. Finalmente, el estudio se centra en la problemática de interés, ya que hayamos ganado la confianza de varios participantes.

Es muy difícil ser “invisibles” en el contexto al momento de ingresar al campo, es ingenuo pensarlo. Pero conforme transcurre el tiempo, los participantes se van acostumbrando a la presencia del investigador y éste va “haciéndose menos visible” (Esterberg, 2002). Por ello, en algunos casos la estancia en el ambiente es larga. Además, quienes no actúan de modo natural, poco a poco van relajándose y su comportamiento resulta cada vez más cotidiano.

Otro aspecto importante es que el investigador nunca debe elevar las expectativas de los participantes más allá de lo posible. A veces las personas piensan que la realización de un estudio implica mejoras en sus condiciones de vida, lo cual no necesariamente es cierto. Entonces tenemos que clarificar que se trata de una investigación cuyos resultados pueden diagnosticar ciertas problemáticas, pero únicamente se limita a esto. Lo más que podemos decir es en dónde

y a quién se les presentará el reporte de resultados. Recordemos: no engañar bajo ninguna circunstancia.

Además, el investigador debe estar abierto a todo tipo de opiniones y escuchar las voces de los diferentes participantes.

Ingresamos al ambiente o campo: ¿y...?

El investigador debe hacer una inmersión total en el ambiente. Regularmente la primer tarea es decidir en qué lugares específicos se recolectarán los datos y quiénes serán los participantes (la muestra). Pero esta labor, a diferencia del proceso cuantitativo, no es secuencial, va ocurriendo y de hecho, la recolección de datos y el análisis ya se iniciaron.

La **inmersión total en el ambiente** implica una serie de actividades físicas y mentales para el investigador o investigadora, entre las que podemos mencionar:

- Observar los eventos que ocurren en el ambiente (desde lo más ordinarios hasta cualquier suceso inusual o importante). Aspectos explícitos e implícitos, sin imponer puntos de vista y tratando, en la medida de lo posible, de evitar el desconcierto o interrupción de actividades de las personas en el contexto. Tal observación es holística (como un “todo” unitario y no en piezas fragmentadas), pero también toma en cuenta la participación de los individuos en su contexto social. El investigador entiende a los participantes, no únicamente registra “hechos” (Williams, Unrau y Grinnell, 2005).
- Establecer vínculos con los participantes, utilizando todas las técnicas de acercamiento (programación neurolingüística, *rappport* y demás que sean útiles), así como las habilidades sociales, de una manera flexible y de acuerdo a las demandas situacionales. Involucrarse con ellos y sus experiencias cotidianas.
- Comenzar a adquirir el punto de vista “interno” de los participantes respecto de cuestiones que se vinculan con el planteamiento del problema. Después podrá tener una perspectiva más analítica o de un observador externo (Williams, Unrau y Grinnell, 2005).
- Recabar datos sobre los conceptos, lenguaje y maneras de expresión, historias y relaciones de los participantes.
- Detectar procesos sociales fundamentales en el ambiente y determinar cómo operan.
- Tomar notas y empezar a generar datos en forma de apuntes, mapas, esquemas, cuadros, diagramas y fotografías, así como recabar objetos y artefactos.
- Elaborar sus descripciones del ambiente (poco más adelante se retomará este punto).
- Estar consciente de su papel y de las alteraciones que provoca (que deben ser anotadas).
- Reflexionar acerca de sus vivencias, que también son una fuente de datos.

Las observaciones durante la inmersión inicial en el campo son múltiples, generales y con poco “enfoque” o dispersas (para entender mejor al sitio y a los participantes o casos). Al principio, el investigador debe observar lo más que pueda. Pero conforme transcurre la investigación, va centrándose en ciertos aspectos de interés (Anastas, 2005; Creswell, 2005; Esterberg, 2002), cada vez más vinculados al planteamiento del problema, que al ser altamente flexible se puede ir modificando. Y el investigador debe prepararse para lo inesperado.

Observación Es necesario ser lo menos obstrutivo que sea posible.

La labor del investigador es como la del detective que arriba a la escena del crimen (de hecho la investigación en Criminalística comienza con la inmersión en el campo): Primero se observa el lugar de forma holística; por ejemplo, si se trata de un asesinato en una casa, se observa toda la habitación donde se encuentra el cadáver (desde las paredes, puertas y ventanas hasta el piso), así como los objetos que hay en el cuarto y el mobiliario. Cada pieza es vista en relación con todo el contexto. Se analiza la posición del cuerpo humano, los gestos de la persona fallecida, los rastros de sangre, etc. Asimismo, se toman muestras de cualquier artefacto o material, desde una posible arma hasta cabellos y fibras de la ropa y del piso, así como rastros de pisadas y huellas. Todo es considerado, y no sólo aquello de la habitación donde se localiza el individuo supuestamente asesinado, sino de cada cuarto y rincón de la casa: jardín, cochera, sótano... Los datos recolectados se envían a un laboratorio para que se les practiquen los análisis apropiados (por ejemplo, tipo de sangre, DNA y composición química). Y conforme la evidencia se va interpretando, el detective enfoca sus observaciones en los elementos vinculados con su problema de investigación: el crimen cometido.

Además, los policías que revisan y evalúan la escena del crimen realizan anotaciones de lo que observan, aun de cuestiones que parecen ser triviales. Si hay datos que no son considerados, se puede perder información valiosa que más adelante podría ser muy útil para responder a las preguntas de investigación: ¿fue realmente un asesinato?, ¿cuándo y cómo ocurrió?, ¿quién pudo ser el asesino?

La mente del investigador al ingresar al campo tiene que ser inquisitiva. De cada observación debe cuestionarse: ¿qué significa esto que observé?, ¿qué me dice en el marco del estudio?, ¿cómo se relaciona con el planteamiento?, ¿qué ocurre o sucedió?, ¿por qué? También es necesario evaluar las observaciones desde diversos ángulos y las perspectivas de distintos participantes (así como el detective visualiza el crimen desde la óptica de la víctima y el asesino, en un estudio sobre la violencia dentro de la familia, la visión de cada miembro es importante).

La **descripción del ambiente** es una interpretación detallada de casos, seres vivos, personas, objetos, lugares específicos y eventos del contexto, y debe transportar al lector al sitio de la investigación (Creswell, 2005). A continuación mostramos un ejemplo de la descripción de un contexto.

EJEMPLO

La iglesia o parroquia de San Juan Chamula, Chiapas, México

A poco más de 10 kilómetros de San Cristóbal de las Casas, en la zona denominada los Altos de Chiapas, en México, se encuentra la comunidad de San Juan Chamula. De manera superficial, parece cualquier pueblo de montaña, pero su organización social y cultura son tan distintas a lo que conocemos que resulta indispensable mantener la mente abierta para descubrirlo. En la plaza central se erige la iglesia de San Juan Chamula (en honor a San Juan Bautista), un hermoso templo edificado en el siglo XVIII. Dicha plaza es una explanada donde se localizan una veintena de puestos en los cuales se venden artesanías (collares, aretes, pulseras, anillos...) así como atuendos para vestir que son confeccionados en tejidos multicolores. En el centro de la plaza está un pequeño kiosco con techo de teja rojiza y columnas verde claro.

La iglesia (también con un techo de teja) se levanta al terminar la plaza, la fachada tiene poco más de 15 metros de altura sobre el piso y hasta el final de su campanario, que incluye tres campanas medianas (no más de un metro de altura) y una cruz en el punto más alto. En general, la edificación es blanca y plana por los costados (salvo un relieve lateral que es un anexo a la parroquia), su portón es de madera y éste, a su vez, tiene en el extremo derecho una puerta más pequeña para ingresar al templo: Alrededor del portón hay un arco pintado en verde claro azulado, que ocupa aproximadamente la tercera parte de todo el edificio y que tiene una ornamentación de cuadrados y rectángulos de no más de 50 centímetros por cada lado con dibujos en relieve de flores, círculos y figuras parecidas a las "X" o taches (verdes, azules, blancos y amarillos). Encima del portal hay otro arco que tiene un balcón, este último arco más pequeño (al igual que el mayor que rodea al portón) tiene los cuadrados multicolores. Además, en los costados de los arcos hay cuatro nichos en colores azul y verde claros.

Por dentro, la iglesia es impresionante: no hay bancos ni bancas ni púlpito, y uno puede observar en el centro del altar a San Juan Bautista (Dios Sol), no a Jesucristo. El piso es de baldosa y el suelo está alfombrado por agujas de pino (éstas forman un pasto seco para "espantar a los malos espíritus"). En las paredes hay algunos troncos de pino recargados. Alrededor del interior de la iglesia se presentan varias figuras de santos, entre ellos: San Agustín, San Pedro y San José. Como los chamulas (indígenas tzotziles que habitan la comunidad), se encomiendan a ellos, los santos "no alcanzan para toda la población". Por eso cada uno fue desdoblado en mayor y menor. Así tenemos entonces un San José Mayor y un San José Menor. Las figuras de los santos llevan colgadas del cuello un espejo y en ocasiones dan la impresión de ser obesos por los muchos vestidos que les van poniendo los fieles que les piden favores. Enfrente (y a veces a un costado) de cada santo hay decenas de velas encendidas colocadas en el piso, lo que hace que en el interior del templo se cuenten por cientos (que cumplen también la función de solicitar favores a los santos, principalmente en cuestiones de salud y bienestar) y que junto con el incienso provocan que el aire esté impregnado de humo y olor. La impresión es mágica y mística. "Al santo que no cumple los rezos le quitan las velas y las colocan a quienes sí cumplen, para que los incumplidos miren cómo se incrementan las velas de sus colegas".

La Virgen María es la Diosa Luna. Está ataviada con prendas multicolores y es una figura hermosa y cautivadora. Por debajo del techo de la iglesia se aprecian unas cuantas mantas de colores más sobrios (de ancho no mayor a un metro) que cuelgan y cruzan de pared a pared (a los costados del templo), parecen bajar del techo de cada lado hasta la mitad de la pared. Como si la iglesia fuera una gran tienda de un sultán en el desierto. En una ocasión se observaron tres mantas y en otra cinco.

En todo el interior del templo se puede escuchar el murmullo por el continuo rezar de los indígenas, donde unos empiezan antes que los otros terminen; de manera tal, que esa especie de "jommm, ummm!", se oye como un sonido grave y profundo permanente, sin interrupciones. Asimismo, en el piso, junto a los santos, se entregan las ofrendas: huevos frescos, gallinas (que son sacrificadas vivas allí mismo), aguardiente y refrescos, en especial los de cola, que sirven para eructar y expulsar a los "malos espíritus".

Los tzotziles (muchos de ellos vestidos en blanco y negro) beben aguardiente en botellas de cristal, sentados en el piso del templo. Algunos rezan solos, otros en grupos pequeños, en ocasiones acompañan sus oraciones con música de guitarra y cantos. Hay quienes, por el

exceso de alcohol están acostados en el piso y completamente embriagados. Los chamulas participan en rituales sincréticos con una devoción y solemnidad única, dialogan con los santos, los increpan, les agradecen, les recriminan, todo de viva voz y en su antigua lengua: el tzotzil. Los chamanes (brujos) rezan y alejan a los malos espíritus, mezclando ritos católicos y paganos.

En el interior, está prohibido tomar fotografías, ya que se corre el peligro de ser agredido y enviado a la cárcel por este hecho, pues los chamulas creen que de esta manera les están robando algo de su alma. Algunos turistas ignorantes de esta advertencia lo han intentado y cuentan que les destrozaron la cámara, los apalearon y enviaron a la cárcel. Fuera de la iglesia, una cruz maya señala los puntos cardinales. Es el árbol de la vida. Al salir, decenas de niños se acercan para vender mercancía, son pobres y comienzan a beber aguardiente prácticamente en los primeros años de su vida.

Los chamulas desterraron de sus templos a los sacerdotes católicos y los convirtieron en recintos con su propia cosmogonía. Las escasas misas se celebran en tzotzil. Tres pinos juntos forman una tríada sagrada que les permite, según su religión, entrar en el más allá. Este interesante concepto tiene gran similitud con el de algunos aborígenes australianos que utilizan los árboles para comunicarse, según ellos, “los de acá” con “los del más allá”. Por esta razón los pinos son parte importante del interior de la iglesia de San Juan de Chamula.²

En el contexto anterior, podría investigarse la religiosidad de los chamulas o sus percepciones sobre el mundo.

También, en la *inmersión inicial* se pueden utilizar diversas herramientas para recabar datos sobre el contexto y completar las descripciones. Por ejemplo, entrevistas y revisión de documentos, que se revisarán en el capítulo 14, relativo a la recolección de los datos. Así pues, recalquemos que toda observación se enmarca.

Es importante completar las descripciones con mapas y fotografías. En el caso de San Juan Chamula esto no es permitido. Algunos lo han hecho, violando el código tzotzil, lo que a nuestro juicio es engañar y un asunto poco ético. El investigador cualitativo en sus observaciones no debe transgredir el ámbito de lo privado y en cambio tiene que ser respetuoso.

No hay un modelo de descripción, cada quien capturará los elementos que le llaman más la atención y esto constituye un dato (como toda la intervención del investigador).

Al estudiante, le recomendamos comenzar con descripciones básicas de objetos (por ejemplo, un jarrón o una bolsa de mano). Luego, descripciones de ambientes con los que se encuentre familiarizado (su recámara, el vestíbulo de su casa, etc.); y posteriormente, de algunas personas (amigos o amigas).

Por otra parte, el investigador escribe lo que observa, escucha y percibe a través de sus sentidos. Dos elementos son comúnmente utilizados: las anotaciones y la bitácora o diario de campo. Usualmente en esta última se registran las primeras.

² Esta última explicación de los pinos proviene de Juan Carrillo, quien en 1981 visitó San Juan Chamula. La referencia en internet es: <http://www.sexovida.com/arte/chamula.htm>

Las anotaciones o notas de campo

Es muy necesario llevar registros y elaborar anotaciones durante los eventos o sucesos vinculados al planteamiento. De no poder hacerlo, la segunda opción es anotar lo más pronto posible después de los hechos. Y como última opción las anotaciones se producen al terminar cada periodo en el campo (al momento de un receso, una mañana o un día, como máximo).

Es conveniente que tales registros y notas se guarden o archiven de manera separada por evento, tema o periodo. Así, los registros y notas del evento o periodo 1, se archivarán de manera independiente de los registros y notas del evento o periodo 2, y de registros y notas del evento o periodo k . Son como páginas separadas que se refieren a los diferentes sucesos o periodos (por ejemplo, por día: lunes, martes, miércoles, jueves, viernes, sábado y domingo). De cada hecho o periodo se anotan la fecha y hora correspondientes. Esto se hace sin importar el medio de registro (computadora de bolsillo, grabadora de voz o video, papel y lápiz, entre otros).

Los materiales de audio y video deben guardarse, y no destruirse, aun cuando se tengan los registros y las anotaciones escritas. También es conveniente tomar fotografías, elaborar mapas y diagramas sobre el contexto, comunidad o ambiente físico (y en ocasiones sus “movimientos” y los de los participantes observados).

Por otro lado, Grimmell (1997) subraya la importancia de incluir nuestros propios términos, palabras, sentimientos y conductas en las anotaciones. Asimismo, cada vez que sea posible (al menos una vez al día) es necesario volver a leer las notas y los registros y, desde luego, anotar nuevas ideas, comentarios u observaciones (y repetir el proceso constantemente).

Las anotaciones pueden ser de diferentes clases y conviene tenerlas todas en la *observación cualitativa*, durante el ingreso al campo y el desarrollo completo del estudio. Básicamente son cinco tipos de anotaciones, que abajo se enumeran.

1. *Anotaciones de la observación directa.* Descripciones de lo que estamos viendo, escuchando, olfateando y palpando del contexto y de los casos o participantes observados. Regularmente van ordenadas de manera cronológica. Nos permitirán contar con una narración de los hechos ocurridos (qué, quién, cómo, cuándo y dónde).

EJEMPLOS

El despertar

Era 10 de noviembre de 2006, 9:30 AM, Andrés entró a la habitación donde estaban reunidos Ricardo y Gabriel; llevaba un conjunto deportivo (*pants*) de color azul marino; su pelo estaba desaliñado; no se había bañado; su mirada reflejaba tristeza; se mostraba cansado. Se sentó en el suelo (en silencio). Ricardo y Gabriel lo observaron y le saludaron con una leve sonrisa; Andrés no respondió. Durante cerca de cinco minutos nadie habló ni miró a los demás. De pronto, Andrés dijo: “Me siento fatal, anoche no debí haber...” Interrumpió su comentario y guardó silencio. Estaba pálido con sus ojos vidriosos y rojos, la boca seca. Se levantó y salió de la habitación, regresó y se limitó a decir: “Voy a la farmacia”, volvió a salir[...]

Testimonio de la Guerra Cristera en México (1926-1929)

Dos jóvenes, R. Melgarejo y Joaquín Silva Córdoba, fueron asesinados en Zamora, Michoacán, el 17 de octubre de 1927. Melgarejo fue obligado a gritar "¡Viva Calles!" En lugar de eso gritó: "¡Viva Cristo Rey!" Entonces, los soldados comenzaron a cortarle las orejas y, al no obtener mejores resultados, le cortaron la lengua. El joven Silva lo abrazó y los soldados les dispararon a ambos, asesinando a los dos jóvenes.³

-
2. *Anotaciones interpretativas.* Comentarios sobre los hechos, es decir, nuestras interpretaciones de lo que estamos percibiendo (sobre significados, emociones, reacciones, interacciones de los participantes). Del tipo: "El director rechaza la tecnología por su edad..."
-

EJEMPLOS

El despertar

Andrés había consumido droga no medicinal la noche anterior y sufría del efecto posterior a tal hecho; probablemente cocaína, que es la sustancia que se acostumbra consumir en este grupo, de acuerdo con observaciones previas. Su salud está muy deteriorada. Incluso, uno de estos días puede morir de una dosis excesiva.

Diario del Che Guevara (7 de octubre de 1967)

Se cumplieron los 11 meses de nuestra inauguración guerrillera sin complicaciones, bucólicamente; hasta las 12.30 hora en que una vieja, pastoreando sus chivas entró en el cañón en que habíamos acampado y hubo que apresarla. La mujer no ha dado ninguna noticia fidedigna sobre los soldados, contestando a todo que no sabe, que hace tiempo que no va por allí. Sólo dio información sobre los caminos; de resultados del informe de la vieja se desprende que estamos aproximadamente a una legua de Higueras, y a otra de Jagüey y a unas dos de Pucará. A las 17.30, Inti, Aniceto y Pablito fueron a casa de la vieja que tiene una hija postrada y una medio enana; se le dieron 50 pesos con el encargo de que no fuera a decir ni una palabra, pero con pocas esperanzas de que cumpla a pesar de sus promesas.

Salimos los 17 con una luna muy pequeña. La marcha fue muy fatigosa y dejando mucho rastro por el cañón donde estábamos, que no tiene casas cerca, pero sí sembradios de papa regados por acequias del mismo arroyo. A las dos paramos a descansar, pues ya era inútil seguir avanzando. El Chino se convierte en una verdadera carga cuando hay que caminar de noche.

El Ejército dio una rara información sobre la presencia de 250 hombres en Serrano para impedir el paso de los cercados en número de 37, dando la zona de nuestro refugio entre el río Acero y el Oro.

La noticia parece diversionista.⁴

³ Parsons (2005, capítulo VIII).

⁴ Ernesto Che Guevara (1967). Referencia de 2005. Aunque no es una anotación interpretativa de una investigación, sí refleja la interpretación de hechos. Esta anotación es lo último que escribió este gran personaje histórico en su diario personal.

3. *Anotaciones temáticas.* Ideas, hipótesis, preguntas de investigación, especulaciones vinculadas con la teoría, conclusiones preliminares y descubrimientos que, a nuestro juicio, vayan arrojando las observaciones. Como: “los habitantes de este municipio no aceptarían a una mujer como alcalde”.

EJEMPLOS

El despertar

“Después de un severo consumo de drogas, al día siguiente los jóvenes de este barrio evitan la comunicación con sus amigos. Las drogas pueden provocar aislamiento”.

La Guerra Cristera en México (1926-1929)

Después de revisar algunos testimonios, puede considerarse que en la Guerra Cristera muchos bandoleros, haciéndose pasar por cristeros, cometieron actos deplorables como saqueos, robos, asesinatos y violaciones a las mujeres. Las guerras civiles son aprovechadas por individuos que en realidad no luchan por un ideal, sino que se aprovechan del caos y la entropía generada.

Experiencias de abuso sexual infantil

Dos tipos de condiciones causales parecen emerger de los datos, las cuales nos conducen a ciertas experiencias fenomenológicas vinculadas al abuso sexual infantil. Estas condiciones pueden ser: a) las normas culturales y b) las formas del abuso sexual. Las normas culturales de dominación y sumisión, la violencia, el maltrato a la mujer, la negación del abuso y la falta de poder de la niña, forman la piedra angular en la cual se perpetra el abuso sexual”.⁵

-
4. *Anotaciones personales* (del aprendizaje, los sentimientos, las sensaciones del propio observador o investigador). Tales como: “La devoción de las Hermanas de María Inmaculada de Guadalupe me resulta admirable”, “El que los empleados sean maltratados va en contra de mis principios”.
-

EJEMPLOS

El despertar

“Me siento triste por Andrés. Me duele verlo así. Está lloviendo y quisiera salirme de la habitación e ir a descansar. Ver tantos problemas me abruma.”

La Guerra Cristera en México (1926-1929)

Cada vez que alguien es perseguido por sus creencias, me parece una injusticia. Y resulta increíble que los primeros cristianos hayan sido perseguidos por la estructura formal del

⁵ Adaptado de Morrow y Smith (1995, p. 6).

Imperio Romano, y en un país latinoamericano, siglos después, se repita la historia. Y en otras épocas, la de la Santa Inquisición, los católicos hayan, esta vez, encarcelado y torturado a quienes no comparten sus creencias. Los extremistas de uno u otro bando, siempre son los que persiguen a quienes no comulgan con su pensamiento. Los extremos son malos.

-
5. *Anotaciones de la reactividad de los participantes* (cambios inducidos por el investigador), problemas en el campo y situaciones inesperadas.
-

EJEMPLO

La violencia intrafamiliar

Al comenzar a entrevistar a las mujeres que parecen ser agredidas por sus esposos, éstos formaron un grupo que fue a hablar con funcionarios de la Alcaldía para protestar por el estudio y presionar nuestra salida.

Estas anotaciones pueden llevarse en una misma hoja en columnas diferentes o vaciarse en páginas distintas; lo importante es que, si se refieren al mismo episodio, estén juntas y se acompañen de las ayudas visuales (mapas, fotografías, videos y otros materiales) e indicaciones pertinentes.

Después, podemos clasificar el material por fecha, temas (por ejemplo, expresiones depresivas, de aliento, de agotamiento), individuos (Andrés, Gabriel, Ricardo), unidades de análisis o cualquier criterio que consideremos conveniente, de acuerdo con el planteamiento del problema.

Luego, se resumen las anotaciones, teniendo cuidado de no perder información valiosa. Por ejemplo, de las notas producto de la observación directa de un episodio entre un médico y un paciente, resumiríamos como se muestra en la tabla 12.3.⁶

En síntesis, las anotaciones:

- Nos ayudan contra la “mala memoria”.
- Señalan lo importante, aquello que contribuya a interpretar y encontrar significado.
- Contienen las impresiones iniciales y las que tenemos durante la estancia en el campo.
- Documentan la descripción del ambiente, las interacciones y experiencias.

Pero el tomar notas no debe interrumpir el flujo de las acciones. Asimismo, en cuanto a éstas debemos evitar generalizaciones *a priori* y juicios de valor imprecisos que a veces son racistas o desprecian a los participantes. Ejemplos de anotaciones erróneas serían: “El sujeto compró muchísimo” (¿qué significa “muchísimo?”), “El cliente come como un cerdo” (¿qué queremos decir?, además es una expresión ofensiva para quien nos ayuda a evaluar un servicio), “El tipo es un patán”, “Ella es una golfa” (¿lo cual significa...?).

⁶ Se ha juntado la letra por cuestiones de espacio.

Tabla 12.3 Un ejemplo de anotaciones resumidas

Resumen	Anotación
El paciente fue sumamente hostil con el médico, tanto verbal como no verbalmente.	Eran las 14:30 horas, cuando en la recepción del hospital, el médico que estaba uniformado con bata blanca, le pidió al paciente que por favor pasara a la sala de espera, con el fin de que se alistara para el chequeo de rutina (su tono fue amable y su comunicación no verbal, afable; miró al paciente directamente a los ojos). El paciente le gritó al médico, con firmeza:
El médico respondió con la misma hostilidad, verbal y no verbal.	"No voy a pasar, váyase a la mierda", y golpeó la pared. No hizo contacto visual con el médico.
Se inició una escalada de violencia verbal.	El médico le respondió al paciente (que por cierto vestía informal): "El que se va a la mierda es usted, púdrase en el infierno" y lanzó el expediente al suelo.
El paciente evadió la interacción.	El paciente contestó: "Mire, matasanos de cuarta categoría, últimamente no me ha dado nada, ni ayudado en nada. Se olvida de los pacientes. No dudo que también lo haga con sus amigos. Ojalá se muera..."
	El paciente visiblemente molesto salió de la recepción del hospital hacia la calle.

Al transcribir las anotaciones (lo cual ocurre tarde o temprano, aunque se trate de grabaciones de audio o video) es conveniente que el interlineado sea doble y con márgenes amplios, tanto por claridad como para facilitar posteriores anotaciones, si después queremos agregar algo.

Anotaciones Se recomienda utilizar oraciones completas.

La bitácora o diario de campo

Asimismo, es común que las anotaciones se registren en lo que se denomina **diario de campo** o **bitácora**, que es una especie de diario personal, donde además se incluyen:

1. Las descripciones del ambiente o contexto (iniciales y posteriores). Recordemos que se describen lugares y participantes, relaciones y eventos, todo lo que juzguemos relevante para el planteamiento.
2. Mapas (del contexto en general y de lugares específicos).
3. Diagramas, cuadros y esquemas (secuencias de hechos o cronología de sucesos, vinculaciones entre conceptos del planteamiento, redes de personas, organigramas, etc.). Tomemos como ejemplo las explosiones de Celaya en septiembre de 1999. Los elementos gráficos se muestran en la figura 12.4.

A las 10:45 horas ocurrió una explosión en una bodega en la que se elaboran y expenden juegos pirotécnicos, la cual sorprendió a comensales de los restaurantes aledaños, transeúntes y comerciantes que se encontraban en la zona de la Central de Abastos, frente a la Central Camionera de Celaya.

Minutos después llegaron bomberos y elementos de la Cruz Roja para tratar de sofocar el incendio y auxiliar a los lesionados. Los cuerpos sin vida empezaban a ser retirados por los paramédicos, mientras los socorristas removían los escombros en busca de otras víctimas, al tiempo que se agrupaban los primeros curiosos, los cuales también fueron sorprendidos poco después.

A las 11:00 horas, las llamas aún fuera de control alcanzaron a tanques de gas estacionario provocando dos explosiones más que atraparon a socorristas y bomberos que cumplían con sus labores de rescate, causando dos bajas entre el grupo de "apagafuegos" y tres en el de paramédicos. Además, cobraron nuevas víctimas entre los lesionados que no pudieron ser evacuados a tiempo. También los primeros informadores en llegar al lugar de los hechos sufrieron bajas. Un fotógrafo de *El Sol del Bajío* perdió la vida. Las cifras preliminares contaban ya 52 cadáveres: 32 varones adultos, 10 mujeres y 10 menores.

A las 14:00 hrs., más de tres horas después de la primera detonación, fue cuando el gobernador del estado de Guanajuato, dio a conocer que la situación estaba "bajo control," y mediante un comunicado de prensa proporcionó el primer reporte oficial, el cual corroboraba las cifras proporcionadas por los rescatistas: 50 muertos y 76 heridos, pero no explicó la existencia de un depósito ilegal de pólvora y materiales inflamables tan grande en pleno centro de la ciudad.

A las 11:40, elementos de la XVI Zona Militar acordonaron una zona que comprende unas 15 calles a partir de la Central de Abastos, incluyendo una terminal de autobuses foráneos, un tianguis, restaurantes y varias tiendas de autoservicio que fueron arrasadas hasta sus cimientos. Se informó que en la zona habitacional más cercana (a 700 metros del siniestro), sólo se registraron algunos vidrios rotos, pero nada de gravedad.

Causa de las explosiones

Corrupción de funcionarios locales y de nivel nacional que permitieron el almacenamiento ilegal de pólvora y otros explosivos.

Almacenamiento ilegal de pólvora y otros explosivos.

Origen de la explosión (desconocido).
No se sabe si fue una chispa, un incendio provocado u otro factor.

Figura 12.4 Explosiones en Celaya (26 de septiembre de 1999)⁷. Cronología de las explosiones.

4. Listados de objetos o artefactos recogidos en el contexto, así como fotografías y videos que fueron tomados (indicando cuándo, fecha y hora, y por qué se recolectaron o grabaron y, desde luego, su significado y contribución al planteamiento).

⁷ Por cuestiones de espacio se resumen los eventos. Basado en relatos de supervivientes y del diario *La Jornada*, 27 de septiembre de 1999, primera plana, columnas 1-3.

EJEMPLO

Guerra Cristera en México, 1926-1929



Fotografía del Claustro de San Francisco en Acámbaro, Michoacán. Se puede observar en los pilares las perforaciones que utilizaban para armar el corral de los caballos del Ejército del Gobierno Mexicano, esto es una muestra de que las iglesias fueron ocupadas y convertidas en cuarteles.

5. Aspectos del desarrollo o curso del estudio (cómo vamos hasta ahora, qué nos falta, qué debemos hacer), que también podrían ser anotaciones temáticas.

EJEMPLO

El descubrimiento de la tumba de Tutankhamon (1922)

Hasta este punto nuestro avance era satisfactorio. Sin embargo, pronto nos dimos cuenta de un hecho más bien preocupante. El segundo féretro que, por lo que se veía a través de la gasa parecía ser una obra de artesanía, presentaba síntomas evidentes del efecto de algún tipo de humedad y en algunos puntos, una tendencia de las incrustaciones a caer. Debo admitir que fue algo desconcertante ya que sugería que había existido algún tipo de humedad antiguamente en el interior de los féretros. De ser así, el estado de conservación de la momia del rey sería menos satisfactorio de lo que habíamos esperado. (Howard Carter)⁸

Como resultado de la inmersión, el investigador debe: *a*) identificar qué tipos de datos deben recolectarse, *b*) en quién o quiénes (muestra), *c*) cuándo (una aproximación) y dónde (lugares específicos, por ejemplo, en una empresa detectar los sitios donde los empleados se reúnen para comentar sus problemas), y *d*) por cuánto tiempo (tentativamente) (Creswell, 2005). Así como definir su papel.

⁸ Carter (1989, p. 173).

Algunas de las actividades que puede realizar un investigador durante la inmersión inicial y el comienzo de la recolección de los datos, son las que se incluyen en la tabla 12.3, algunas de las cuales ya se comentaron y otras se revisarán en los siguientes capítulos.

La inmersión en el contexto, ambiente o campo Implica, a veces, vivir en éste o ser parte de él (trabajar en la empresa, habitar en la comunidad, etcétera).

Tabla 12.4 Cuestiones importantes en el trabajo de campo de una investigación cualitativa

Acceso al contexto, ambiente o sitio

- Elegir el contexto, ambiente o sitio.
- Evaluar nuestros vínculos con el contexto.
- Lograr el acceso al contexto o sitio y a los participantes.
- Contactar a las personas que controlan la entrada al ambiente o sitio y tienen acceso a los lugares y personas que lo conforman (*gatekeepers*), así como obtener su buena voluntad y participación.
- Realizar una inmersión completa en el contexto y evaluar si es el adecuado de acuerdo con nuestro planteamiento.
- Lograr que los participantes respondan a las solicitudes de información y aporten datos.
- Decidir en qué lugares del contexto se recolectan los datos.
- Planear qué tipos de datos se habrán de recolectar.
- Desarrollar los instrumentos para recolectar los datos (guías de entrevista, guías de observación, etcétera).

Observaciones

- Registrar notas de campo creíbles, desde el ingreso al ambiente (impresiones iniciales) hasta la salida; escritas o grabadas en algún medio electrónico.
- Registrar citas textuales de los participantes.
- Definir y asumir el papel de observador.
- Transitar en la observación: Enfocar paulatinamente de lo general a lo particular.
- Validar si los medios planeados para recolectar los datos son las mejores opciones para obtener información.

Entrevistas iniciales

- Planearlas cuidadosamente.
- Concertar citas.
- Preparar el equipo para grabar las entrevistas.
- Acudir a las citas puntualmente.
- Realizar las entrevistas.
- Registrar anotaciones y hechos relevantes de las entrevistas.

Documentos

- Elaborar listas de lugares donde se pueden localizar y obtener documentos.
- Tramitar los permisos para obtenerlos o reproducirlos.
- Preparar el equipo para escanear, videograbar o fotografiar los documentos.
- Cuestionar el valor de los documentos.
- Certificar la autenticidad de los documentos.

Bitácora y diarios

(continúa)

Tabla 12.4 Cuestiones importantes en el trabajo de campo de una investigación cualitativa (continuación)

- Solicitar a los participantes que escriban diarios y bitácoras.
- Revisar periódicamente esos diarios y bitácoras.

Materiales y objetos

- Recolectar, grabar o tomar videos, fotografías, audiocintas y todo tipo de objetos o artefactos que puedan ser útiles.

Como podemos ver, las fases del proceso investigativo se traslapan y no son secuenciales, uno puede regresar a una etapa inicial y retomar otra dirección. El planteamiento puede variar y tomar rumbos que ni siquiera habíamos previsto. Por ejemplo, en la investigación sobre las emociones que pueden experimentar los pacientes jóvenes que tendrán una intervención quirúrgica por un tumor, podemos concentrarnos en el estrés o en temores específicos; o bien, enfocarnos en el trato que reciben por parte de los médicos, enfermeras y personal auxiliar que los atiende antes de la operación. Podemos también estudiar sus emociones antes y después de la operación, incluir o no a sus familiares; en fin, el laberinto puede conducirnos a varias partes.

La inmersión inicial (que a estas alturas envuelve el trabajo para la recolección de los datos), nos conduce a seleccionar un diseño y una muestra que, como veremos, implica continuar adentrándonos en el ambiente y tomar decisiones.

- Los planteamientos cualitativos suelen incluir los objetivos, las preguntas de investigación, la justificación y la viabilidad, además de una exploración de las deficiencias en el conocimiento del problema y la definición inicial del ambiente o contexto.
- Los objetivos y preguntas son más generales y enunciativos en los estudios cualitativos.
- Los elementos de justificación en los planteamientos cualitativos son los mismos que en los cuantitativos: conveniencia, relevancia social, implicaciones prácticas, valor teórico y utilidad metodológica.
- La flexibilidad de los planteamientos cualitativos es mayor que la de los cuantitativos.
- Los planteamientos cualitativos son: abiertos, expansivos, no direccionados en su inicio, fundamentados en la experiencia e intuición, se aplican a un número pequeño de casos, el entendimiento del fenómeno es en todas sus dimensiones, se orientan a aprender de experiencias y puntos de vista de los individuos, valorar procesos y generar teoría fundamentada en las perspectivas de los participantes.
- Para responder a las preguntas de investigación es necesario elegir un contexto o ambiente donde se lleve a cabo el estudio; asimismo, es preciso ubicar el planteamiento en espacio y tiempo.



- Algunos autores consideran que para plantear el problema no se debe efectuar una revisión de la literatura; para otros sí, pero en tal caso, el papel de ésta es únicamente de apoyo. La investigación cualitativa se basa, ante todo, en el proceso mismo de recolección y análisis de los datos.
- Para quienes se inician en la investigación cualitativa, se sugiere visualizar gráficamente el problema de estudio.
- Los planteamientos cualitativos son una especie de plan de exploración y resultan apropiados cuando el investigador se interesa por el significado de las experiencias y valores humanos, el punto de vista interno e individual de las personas y el ambiente natural en que ocurre el fenómeno estudiado; así como cuando buscamos una perspectiva cercana de los participantes.
- En los estudios cualitativos, las hipótesis adquieren un papel distinto al que tienen en la investigación cuantitativa. Normalmente no se establecen antes de ingresar en el ambiente y comenzar la recolección de los datos. Más bien, durante el proceso, el investigador va generando hipótesis de trabajo que se afinan paulatinamente conforme se recaban más datos o las hipótesis son uno de los resultados del estudio.
- Ya que se ha elegido un ambiente o lugar apropiado, comienza la tarea de responder a las preguntas de investigación. El ambiente puede ser tan variado como el planteamiento del problema.
- Tal ambiente puede variar, ampliarse o reducirse y es explorado para ver si es el apropiado.
- Dos dimensiones resultan esenciales con respecto a la selección del ambiente: conveniencia y accesibilidad.
- Para lograr el acceso al ambiente debemos negociar con los *gatekeepers*.
- Con el fin de tener un mayor y mejor acceso al ambiente, así como ser aceptados, se recomienda: desarrollar relaciones, elaborar una historia sobre la investigación, no intentar imitar a los participantes, planear el ingreso y no elevar las expectativas más allá de lo necesario.
- La inmersión total implica observar eventos, establecer vínculos con los participantes, comenzar a adquirir su punto de vista; recabar datos sobre sus conceptos, lenguaje y maneras de expresión, historias y relaciones; detectar procesos sociales fundamentales. Tomar notas y empezar a generar datos en forma de apuntes, mapas, esquemas, cuadros, diagramas y fotografías, así como recabar objetos y artefactos; elaborar descripciones del ambiente. Estar consciente del propio papel como investigador y de las alteraciones que se provocan; así como reflexionar acerca de las vivencias.
- Las observaciones al principio son generales pero van enfocándose en el planteamiento.
- La descripción del ambiente es una interpretación detallada de casos, seres vivos, personas, objetos, lugares específicos y eventos del contexto, y debe transportar al lector al sitio de la investigación.
- Se deben tomar distintos tipos de anotaciones: de la observación directa, interpretativas, temáticas, personales y de reactividad de los participantes.

- Las anotaciones se registran en el diario o bitácora de campo, que además contiene: descripciones, mapas, diagramas, esquemas, listados y aspectos del curso del estudio.
- Para complementar las observaciones podemos realizar entrevistas, recolectar documentos, etcétera.

CONCEPTOS BÁSICOS

Acceso al contexto o ambiente

Ambiente (contexto)

Anotaciones de campo

Bitácora de campo (diario de campo)

Descripciones del ambiente

Gatekeepers

Hipótesis de trabajo

Inmersión inicial en el campo

Inmersión total en el campo

Justificación del estudio

Literatura

Objetivos de investigación

Observación

Participantes

Planteamiento del problema

Preguntas de investigación

Proceso cualitativo

Viabilidad del estudio



EJERCICIOS

1. Vea la película de moda y plantee un problema de investigación cualitativa (como mínimo objetivos, preguntas y justificación de la investigación). ¿Cuál es el contexto o ambiente inicial de dicho estudio?
2. Seleccione un artículo de una revista científica que contenga los resultados de una investigación cualitativa y responda las siguientes preguntas: ¿cuáles son los objetivos de esa investigación?, ¿cuáles son las preguntas?, ¿cuál es su justificación?, ¿cuál su contexto o ambiente?, ¿cómo el (los) investigador(es) ingresó(aron) al campo?
3. Visite una comunidad rural y observe qué sucede en tal comunidad, platique con sus habitantes y recolecte información sobre un asunto que le interese. Tome notas y analícelas. De esta experiencia, plantee un problema de investigación cualitativo.
4. Respecto de la idea que eligió en el capítulo 2 y que fue desarrollando bajo el enfoque cuantitativo a lo largo del libro, ahora transfórmela en un planteamiento de investigación cualitativo. ¿Cuál sería el contexto o ambiente inicial de este planteamiento?
5. Describa un cuadro (pintura) de un artista renacentista, uno de un impresionista, otro de un cubista, uno más de un surrealista y finalmente de un pintor del siglo XXI. Analice y compare sus descripciones.
6. Vamos a presentar distintas anotaciones y le pediríamos que nos indicara de qué tipo son en cada caso. Las respuestas se incluyen en la sección del CD anexo: Respuestas a los ejercicios.

Anotación	Tipo de anotación a que corresponde
"El día 10 de noviembre, los diputados del partido no asistieron al Comité de Seguridad Nacional..."	
"La relación de este matrimonio es asimétrica porque la autoestima de él es muy baja..."	
"No habrá acuerdo parlamentario para aprobar la reforma energética, debido a que éste no conviene a los intereses de los políticos del partido..."	
"El alcoholismo en la comunidad se está fomentando mediante la apertura de bares y tabernas en la zona sur".	
"Los padres de familia se están negando a que se realicen las entrevistas con sus hijos".	
"Tales diputados son unos incompetentes, su ignorancia me hace rechazarlos".	
"Me provoca náuseas que los hombres agreden de esta forma a sus esposas..."	
"Los contenedores de las piezas recibidas en la terminal marítima 28, el día 23 de agosto, a las 21:00 horas, fueron colocados en desorden..."	
"La causa del conflicto en la empresa RTC-DEL es la rivalidad entre el líder sindical Fernando y la directora Abigail..."	

7. Realice una inmersión inicial en el campo de una fábrica, una escuela, un hospital, un barrio, una fiesta o una comunidad, teniendo en mente un planteamiento. ¿Quiénes son los *gatekeepers*?, ¿quiénes son los participantes?,

¿qué acontecimientos le llamaron la atención?, ¿qué datos pueden recolectarse y ser útiles para el estudio planteado? Describa un lugar específico del ambiente o contexto.

EJEMPLOS DESARROLLADOS

LA GUERRA CRISTERA EN GUANAJUATO

Una breve explicación

La "Cristiada" se le denominó a la guerra que, entre 1926 y 1929, enfrentó –en México– al Gobierno con la Iglesia católica. Las relaciones entre ambos poderes eran conflictivas de años atrás y se politizaron con la división de liberales y conservadores durante el conflicto armado, que fue en realidad una guerra civil. Mientras la Iglesia apoyó a los conservadores y propuso la cristiandad como solución, los liberales abogaban por la secularización de los bienes del clero y la abolición de las órdenes religiosas (Scavino, 2005).

Los antecedentes son diversos y comienzan desde principios del siglo XIX. Pero desembocaron en varios acontecimientos que es necesario destacar:

- En 1924, el general Plutarco Elías Calles asume la Presidencia de México.
- El 21 de febrero de 1925, los caudillos de la Confederación Regional Obrera Mexicana (CROM), auspicados por Calles, proclaman el surgimiento de la "Iglesia católica apostólica mexicana" (Carrère, 2005). Esto implicaba una especie de ruptura con la autoridad del Vaticano.
- El proyecto de tal institución fracasó rotundamente. El Papa Pío XI en la encíclica *Quas Primas*, del 11 de diciembre de 1925, declara de manera universal la festividad de Cristo Rey (Carrère, 2005).
- En el centro de México, Cristo Rey era un símbolo fundamental del catolicismo; incluso, años atrás habían erigido un monumento en el estado de Guanajuato a tal figura.
- El 2 de febrero de 1926, Pío XI dirige al Episcopado mexicano una carta en la cual insta a los católicos a emprender la acción cívica contra algunas medidas persecutorias que comenzaban a materializarse, pero que se abstuvieran de

formar un partido político, para así evitar acusaciones por parte del gobierno de Calles de intervenir en asuntos políticos (Carrère, 2005).

- En marzo, se crea la Liga Nacional de la Defensa de la Libertad Religiosa, la cual defendería los derechos de profesar, confesar y promover la "Fe católica" (Carrère, 2005).

Plutarco Elías Calles, el presidente mexicano más radical en materia religiosa, obtuvo del Congreso, en enero de 1926, la aprobación de la Ley Reglamentaria del artículo 130, la cual facultaba al Poder Federal la regulación de la "disciplina" de la Iglesia y confirmaba el desconocimiento de la personalidad jurídica de la Iglesia, de tal suerte que los sacerdotes serían considerados como simples profesionistas y las legislaturas estatales tendrían facultad para determinar el número máximo de sacerdotes dentro de su jurisdicción. Se requería, además, un permiso del Ministerio del Interior (actualmente Secretaría de Gobernación) para la apertura de nuevos lugares de culto

(Dirección General del Archivo Histórico del Senado, 2003). Los sacerdotes debían registrarse ante el Ministerio del Interior.

- El 31 de julio de 1926 entra en vigor la "Ley Calles". En tanto, el Episcopado mexicano consulta con el Vaticano en Roma para suspender los cultos en las iglesias el mismo 31 de julio. El Papa aprueba las medidas propuestas por la Iglesia mexicana. El general Calles, al conocer las intenciones de los católicos, ordena que las iglesias sean cerradas e inventariadas (Dirección General del Archivo Histórico del Senado, 2003).
- El cierre de templos origina una gran cantidad de protestas oficiales de la Iglesia mexicana, y la Ley Calles, en la práctica, se convierte en acciones tales como la prohibición del culto religioso, del suministro de sacramentos, de la catequesis; la supresión de monasterios y conventos, la libertad de prensa religiosa y la expropiación de algunos