

CAPÍTULO

7

RECOLECCIÓN DE LOS DATOS CUANTITATIVOS: SEGUNDA PARTE

INTRODUCCIÓN

En este capítulo se tratan algunos temas que por cuestiones de espacio no fueron incluidos en el capítulo 9 (“Recolección de los datos cuantitativos”) del texto impreso de *Metodología de la investigación*, 5ª edición; fundamentalmente los siguientes: procedimiento “paso a paso” para construir un instrumento de medición, el escalograma de Guttman (escala de actitudes), pruebas e inventarios, análisis de contenido, observación cuantitativa, y obtención y análisis de datos secundarios (comprendido el metaanálisis) y un comentario sobre las fuentes de error en las mediciones.

EL PROCEDIMIENTO PARA CONSTRUIR UN INSTRUMENTO DE MEDICIÓN “PASO A PASO”

En el capítulo 9 del texto impreso: “Recolección de los datos cuantitativos” (sección: ¿Qué procedimiento se sigue para construir un instrumento de medición?), se comentó en términos generales el proceso para desarrollar un instrumento o sistema de recolección de los datos. Ahora revisaremos con mayor profundidad cada una de las fases.

Fase 1: Redefiniciones fundamentales

Al comenzar a desarrollar el o los instrumentos de medición, vale la pena una última 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)? _____. ¿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 sistema electrónico, etcétera.) _____.

Como resultado de responder 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 pueden sernos 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, Punch, 2009).
- La muestra: ¿a quién o quiénes fue administrado?, ¿qué perfil o características poseen?
- Las condiciones de aplicación: ¿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 aplicado, 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 la variable. Incluso, se sugiere hacerlo gráficamente como en la figura 7.1., con la “satisfacción laboral”).



Figura 7.1 Ejemplo con la variable satisfacción laboral: establecimiento de factores

Una vez ubicados tales factores, establecer los indicadores de cada uno. Por ejemplo, en el caso de la “satisfacción en torno a las condiciones de trabajo”, éstos podrían ser: satisfacción sobre la 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.

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 fiabilidad, 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 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 revisan 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 para 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 su 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 reactivos.

El tránsito de variable a ítems es progresivo, como se muestra en el siguiente ejemplo:

EJEMPLO

Variable	Dimensión	Indicadores	Ítems
Intención del voto para las próximas elecciones para Alcalde	Intención de voto por partido	Intención de voto favorable	En las próximas elecciones para Alcalde, ¿por qué partido piensa votar? (con sus opciones de respuesta)
		Intención de voto desfavorable	En las próximas elecciones para Alcalde, ¿por qué partido nunca votaría? (con sus opciones de respuesta)
	Intención de voto por candidato	Intención de voto favorable	En las próximas elecciones para Alcalde, ¿por qué candidato piensa votar? (con sus opciones de respuesta)
		Intención de voto desfavorable	En las próximas elecciones para Alcalde, ¿por qué candidato nunca votaría? (con sus opciones de respuesta)

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, 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 o una página web, e incluso que un cuestionario en una palm o computadora u ordenador portátil.

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, que 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 comentó sobre las formas de aplicar cuestionarios se profundizó en el tema, pero en el caso de pruebas estandarizadas éstas suelen administrarse en lugares que permitan a los participantes concentrarse, como aulas universitarias, salones de hoteles y despachos privados. En cambio, la observación depende de los hechos o conductas que se pretende registrar.

Fase 5: Construcción del instrumento

Esta etapa implica la generación de todos los ítems o reactivos y categorías del instrumento, que 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 (codificación).

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 tentativa 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) aplicó 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 contuvo 99 reactivos.

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 (Neuman, 2009). 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.

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-evitación 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 participantes 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 7.1¹

Tabla 7.1 Recomendaciones para la prueba piloto en instrumentos escritos

- Es aconsejable que se dispongan espacios para comentarios de los 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).

Fase 7: Versión final

Así, con los resultados de la prueba piloto, el instrumento de medición preliminar se modifica, ajusta y mejora (se quitan o agregan ítems, se cambian palabras, se otorga más

¹ Algunas de estas sugerencias fueron adaptadas de Mertens (2005) y Koretz (2009).

tiempo para responder, etc.). Se tiene la versión final para administrar, la cual incluye un diseño gráfico atractivo para los participantes y de acuerdo con el medio (impreso en papel, CD, página web o blog, palm o PC, etcétera).

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 individuos 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 respecto a 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 la 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 de identificación
- k)* El instrumento de medición (estructura, orden, ítems, instrucciones para el personal y para los 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 este CD anexo, se muestra el ejemplo del proyecto de Núñez (2001).

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 (consultar capítulo 2 de este CD).

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. 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, hemos decidido ver en el capítulo 9 del texto impreso y en este capítulo del CD cada instrumento y los contextos más comunes en que se usan.

ESCALOGRAMA DE GUTTMAN

En el libro *Metodología de la investigación*, 5ª edición, concretamente en las escalas de actitudes, como técnicas para recolectar datos, se comentaron el escalamiento tipo Likert y el diferencial semántico, por lo que quedó pendiente el escalograma de Guttman. Ahora lo consideraremos.²

Este método para medir actitudes fue desarrollado por Louis Guttman (1916-1987).³ Se basa en el principio de que algunos ítems indican en mayor medida la fuerza o intensidad de la actitud. *La escala está constituida por afirmaciones*, las cuales poseen las mismas características que el escalamiento tipo Likert. Pero *el escalograma garantiza que la escala mida una dimensión única*. Es decir, cada afirmación mide la misma dimensión de la misma variable. A esta propiedad se le conoce como *unidimensionalidad* (Reckase, 2009).

Algunos autores consideran que el escalograma, más que un método de medición de actitudes, es una técnica para determinar si un conjunto de afirmaciones reúne los requisitos de un tipo particular de escala (Edwards, 1983; Engelhard, 2008).

Para construir el escalograma *es necesario desarrollar un conjunto de afirmaciones pertinentes al objeto de actitud. Éstas deben variar en intensidad* (Trochim, 2008). Por ejemplo, si pretendiéramos medir la actitud hacia la calidad en el trabajo dentro del nivel gerencial, la afirmación: “la calidad tiene que vivirse en todas las actividades del trabajo y en el hogar”, es más intensa que la afirmación: “la calidad debe vivirse sólo en las actividades más importantes del trabajo”. Dichas afirmaciones *se aplican a una muestra a manera de prueba piloto. Y una vez administradas, se procede a su análisis*. Cabe mencionar que *las categorías de respuesta* para las afirmaciones *pueden variar entre dos* (“de acuerdo-en desacuerdo”, “sí-no”, etc.) *o más categorías* (por ejemplo, las mismas categorías que en el caso de Likert).

2 Para profundizar en esta escala, se sugiere consultar a Roberts y Jowell (2008); Dunn-Rankin et al. (2004); Cliff, Kyats y Keats (2002), Weller y Romney (1990), Black y Champion (1976), Guttman (1976), Lingoes (1976), Dotson y Sumers (1976) y Nie et al. (1975).

3 Guttman hizo importantes aportaciones a la medición, además del escalograma, entre otras cuestiones, fue quien conceptualizó en 1944 los niveles de medición nominal, ordinal, de intervalos y de razón (Katz, 1988, Stevens, 1951).

Técnica de Cornell

La manera más conocida de analizar los ítems o las afirmaciones y desarrollar el escalograma es la *técnica de Cornell* (Guttman, 1976). En ella se procede a:

1. *Obtener el puntaje total* de cada participante o sujeto en la escala.
2. *Ordenar a los sujetos o encuestados* de acuerdo con su puntaje total (del puntaje mayor al menor, de manera vertical descendente).
3. *Ordenar las afirmaciones* de acuerdo con su intensidad (de mayor a menor y de izquierda a derecha).
4. *Construir un cuadro donde se crucen los puntajes de los participantes ordenados con los ítems y sus categorías, de manera jerarquizada.* Así, tenemos un cuadro donde los sujetos constituyen los renglones y las categorías de los ítems forman las columnas. Esto se representa en la figura 7.2.
5. *Analizar el número de errores o rupturas en el patrón ideal de intensidad* de la escala.

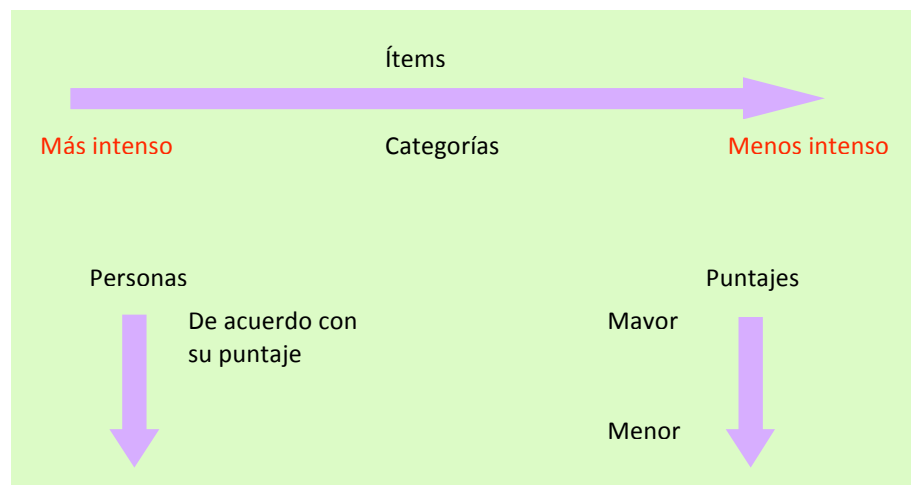


Figura 7.2 Manera de disponer los resultados para el análisis de ítems mediante la técnica de Cornell

EJEMPLO

Suponga que aplicamos una escala con cuatro ítems o afirmaciones a 14 sujetos o participantes, cuyo objeto de actitud es “la calidad dentro de la organización” (la variable es la conciencia sobre la calidad por parte de los empleados).

La calidad en la organización, definida como: “el grado en que los empleados de una empresa se orientan a cumplir con las especificaciones y requerimientos de los clientes en cuanto a sus productos y servicios” (en precio, tiempo de entrega, cumplimiento de estándares del mercado, etcétera).

Cada afirmación tiene dos categorías de respuesta (“de acuerdo” y “en desacuerdo”, codificadas como uno y cero, respectivamente).

Afirmación A: “la calidad tiene que vivirse siempre en todas las actividades del trabajo y en el hogar”.

Afirmación B: “la calidad tiene que vivirse siempre en todas las actividades del trabajo”.

Afirmación C: “la calidad tiene que vivirse solamente en las actividades más importantes del trabajo”.

Afirmación D: “la calidad tiene que vivirse solamente en la actividad más importante del trabajo”.

Los resultados se muestran en la tabla 7.2.

Tabla 7.2 Ejemplo de la técnica de Cornell para el análisis de ítems o afirmaciones

SUJETO	AFIRMACIONES								PUNTUACIONES TOTALES
	A		B		C		D		
	DA (1)	ED (0)	DA (1)	ED (0)	DA (1)	ED (0)	DA (1)	ED (0)	
1	X		X		X		X		4
2	X		X		X		X		4
3	X		X		X		X		4
4		X	X		X		X		3
5		X	X		X		X		3
6		X	X		X		X		3
7		X		X	X		X		2
8		X		X	X		X		2
9		X		X	X		X		2
10		X		X	X		X		2
11		X		X		X	X		1
12		X		X		X	X		1
13		X		X		X		X	0
14		X		X		X		X	0

DA = de acuerdo o 1, ED = en desacuerdo o 0

Ahora bien, como se observa en la tabla 7.2, los participantes están ordenados por su puntuación en la escala total. Asimismo, las frases deben ordenarse por su intensidad (en el ejemplo, *A* tiene mayor intensidad que *B*, *B* mayor que *C*, y a su vez, *C* mayor que *D*) también sus categorías se encuentran jerarquizadas de acuerdo con su valor, de izquierda a derecha. Hay que recordar que si la afirmación es negativa, la codificación se invierte (si para la frase: *la calidad es poco importante para el desarrollo de una empresa*, se tiene la respuesta: “de acuerdo” se codificará con cero, y “en desacuerdo” con uno). En el ejemplo de la tabla 7.2 encontramos cuatro afirmaciones positivas.

Los sujetos que estén “de acuerdo” con la afirmación *A*, que es la más intensa, muy probablemente también lo estén con las afirmaciones *B*, *C* y *D*, ya que su intensidad es menor. Los individuos que respondan “de acuerdo” a la afirmación *B* tenderán a estar “de acuerdo” con *C* y *D* (afirmaciones menos intensas), pero no necesariamente con *A*.

Quienes estén “de acuerdo” con *C* lo más probable es que se encuentren “de acuerdo” con *D*, pero no necesariamente con *A* y *B*.

Debe observarse que el individuo número 1 estuvo “de acuerdo” respecto a las cuatro afirmaciones. Los sujetos 2 y 3 respondieron de igual forma. Las puntuaciones de éstos equivalen a cuatro ($1 + 1 + 1 + 1$). Los participantes 4, 5 y 6 obtuvieron una puntuación de tres, pues estuvieron “de acuerdo” con tres afirmaciones y así sucesivamente (los últimos dos sujetos estuvieron “en desacuerdo” respecto a todas las afirmaciones).

En forma ideal, los sujetos que obtienen una puntuación total de cuatro en esta escala particular, es porque respondieron “de acuerdo” a las cuatro afirmaciones. Los individuos que alcanzan una puntuación total de tres respondieron que están “en desacuerdo” con la primera afirmación, pero “de acuerdo” con las demás afirmaciones. Quienes reciben una puntuación de dos manifiestan estar “en desacuerdo” con los dos primeros ítems, pero “de acuerdo” con los dos últimos. Del mismo modo, los individuos con puntuación de uno respondieron “en desacuerdo” a las tres primeras afirmaciones y “de acuerdo” a la última. Por último, aquellos que estuvieron “en desacuerdo” respecto a las cuatro afirmaciones, recibieron una puntuación total de cero.

Los sujetos o participantes se escalan de manera perfecta, sin que nadie rompa el patrón de intensidad de las afirmaciones: si están “de acuerdo” con la afirmación más intensa, también lo están con las menos intensas. En otro ejemplo, si estoy de acuerdo con la afirmación: *“podría casarme con una persona de nivel económico diferente al mío”*, seguramente estaré de acuerdo con la afirmación: *“podría viajar en un automóvil con una persona de nivel económico diferente al mío”* (casarse es más intenso que viajar).

Cuando los individuos se escalan de manera perfecta respecto a las afirmaciones, esto quiere decir que los ítems en realidad varían gradualmente en intensidad. Es la prueba empírica de que están escalados por su intensidad. Se le denomina *reproductividad* al grado en que un conjunto de afirmaciones o ítems escalan perfectamente según su intensidad. Esto significa que el patrón de respuesta de una persona en relación con todos los ítems puede reproducirse con exactitud tan sólo con conocer su puntuación total en la escala completa (Black y Champion, 1976; Roberts y Jowell, 2008).

La reproductividad ideal se da cuando nadie rompe el patrón de intensidad de la escala. Sin embargo, en la realidad sólo algunas cuantas escalas del tipo de Guttman reúnen la reproductividad ideal; la mayoría contienen inconsistencias o rupturas al patrón de intensidad. El grado en que se alcanza el patrón perfecto de intensidad de la escala, o *reproductividad*, se determina con el análisis del número de personas o casos que rompen dicho patrón, que es el quinto paso para construir el escalograma de Guttman.

Análisis del número de errores o rupturas del patrón ideal de intensidad de la escala

Un error es una inconsistencia en las respuestas de una persona a una escala; es un rompimiento con el patrón ideal de intensidad de la escala. La tabla 7.3 muestra tres ejemplos de error encerrados en círculos que, como puede verse, son inconsistencias del patrón ideal. El segundo participante respondió “de acuerdo” a los ítems más intensos o fuertes y “en desacuerdo” al ítem menos intenso. El cuarto sujeto manifestó estar “de acuerdo” con las afirmaciones *B* y *C* (supuestamente más intensas), pero “en desacuerdo” con *D* (supuestamente menos intensa). El quinto sujeto estuvo “en desacuerdo” con los ítems menos intensos, pero “de acuerdo” con el más intenso. Son inconsistencias o errores. Si un escalograma presenta diversos errores significa que los ítems en realidad no tienen distintos niveles de intensidad (o que los participantes no entienden o son deshonestos).

Tabla 7.3 Ejemplos de errores o inconsistencias en un escalograma de Guttman

AFIRMACIONES									
SUJETO	A		B		C		D		PUNTUACIONES TOTALES
	DA	ED	DA	ED	DA	ED	DA	ED	
1	X		X		X		X		4
2	X		X		X			X	3
3		X	X		X		X		3
4		X	X		X			X	2
5	X			X		X		X	1
6				X		X		X	0

DA = de acuerdo o 1, ED = en desacuerdo o 0

Los *errores* se detectan mediante el análisis de las respuestas que rompen el patrón; para ello, se establecen los *puntos de corte* en el cuadro donde se cruzan las afirmaciones y sus categorías con las puntuaciones totales. En el ejemplo de la tabla 7.2, los “puntos de corte” serían los que se muestran en la tabla 7.4.

Tabla 7.4 Ejemplo de establecimiento de los puntos de corte en la técnica de Cornell

SUJETO	AFIRMACIONES								PUNTUACIONES TOTALES
	A		B		C		D		
	DA (1)	ED (0)	DA (1)	ED (0)	DA (1)	ED (0)	DA (1)	ED (0)	
1	X		X		X		X		4
2	X		X		X		X		4
3	X		X		X		X		4
4		X	X		X		X		3
5		X	X		X		X		3
6		X	X		X		X		3
7		X		X	X		X		2
8		X		X	X		X		2
9		X		X	X		X		2
10		X		X	X		X		2
11		X		X		X	X		1
12		X		X		X	X		1
13		X		X		X		X	0
14		X		X		X		X	0

DA = de acuerdo o 1, ED = en desacuerdo o 0

No se aprecia ninguna inconsistencia. En cambio, en la tabla 7.5 se aprecian tres inconsistencias o errores; las respuestas están desubicadas respecto de los *puntos de corte*, es decir, rompen el patrón de intensidad. (Los errores están encerrados en un círculo.)

Tabla 7.5 Ejemplo de errores respecto a los puntos de corte

AFIRMACIONES									
SUJETO	A		B		C		D		PUNTUACIONES TOTALES
	DA	ED	DA	ED	DA	ED	DA	ED	
	(1)	(0)	(1)	(0)	(1)	(0)	(1)	(0)	
1	X		X		X		X		4
2	X		X		X		X		4
3	X		X		X		X		4
4		X	X		X		X		3
5		X	X		X		X		3
6		X	X		X		X		3
7	X			X	X		X		3
8		X		X	X		X		2
9		X		X	X		X		2
10	X			X	X		X		2
11		X		X		X	X		1
12	X			X		X	X		1
13		X		X		X		X	0
14		X		X		X		X	0

----- = Puntos de corte (líneas punteadas)

Como se mencionó anteriormente, cuando el número de errores es excesivo la escala no presenta *reproductividad* y no puede aceptarse. La *reproductividad* se determina mediante un *coeficiente*. La fórmula de este coeficiente es:

$$\text{Coeficiente de reproductividad} = 1 - \left(\frac{\text{Número de errores o inconsistencias}}{\text{Número total de repuestas}} \right)$$

Donde el número total de respuestas es igual al número de ítems o afirmaciones multiplicado por el número de sujetos. Por tanto, la fórmula directa sería:

$$\text{Coeficiente de reproductibilidad} = 1 - \left(\frac{\text{Número de errores}}{(\text{número de ítems})(\text{número de sujetos})} \right)$$

En el ejemplo de la tabla 7.5, tendríamos que el coeficiente de reproductividad sería:

$$Cr = 1 - \frac{3}{(4)(14)}$$

$$Cr = 1 - 0.053$$

$$Cr = 0.947 = 0.95$$

El coeficiente de reproductividad oscila entre 0 y 1, cuando equivale a 0.90 o más nos indica que el número de errores es tolerable, además de que la escala es unidimensional y se acepta. Cuando es menor que 0.90 no se acepta la escala. Originalmente Guttman recomendó administrar un máximo de 10 a 12 ítems o afirmaciones a un mínimo de 100 personas (Black y Champion, 1976).

Una vez determinado el número de errores aceptable mediante el *coeficiente de reproductividad*, se procede a aplicar la escala definitiva (si dicho coeficiente fue de 0.90 o más; esto es, si el error permitido no excedió 10%) o a realizar ajustes en la escala (reconstruir ítems, eliminar ítems que estén generando errores, etc.). Los cinco pasos mencionados son una especie de prueba piloto para demostrar que la escala es unidimensional y funciona.

Codificación de respuestas

Cuando se aplica la versión definitiva de la escala, los resultados se *codifican* de la misma manera que en la escala Likert, esto depende del número de categorías de respuesta que se incluyan. Al igual que la escala Likert y el diferencial semántico, *todos los ítems deben tener el mismo número de categorías de respuesta. Éste es un requisito de todas las escalas de actitud*. Asimismo, se considera una respuesta inválida si se marcan dos o más opciones para una misma afirmación. *El escalograma de Guttman es una escala estrictamente ordinal, pero que se suele usar como si fuera de intervalo*. Puede aplicarse mediante *entrevista* (con uso de tarjetas que contengan las opciones o categorías de respuesta) o *autoadministrarse*.

PRUEBAS E INVENTARIOS

En las ciencias y disciplinas del siglo XXI gran cantidad de estudios utilizan para sus mediciones pruebas e inventarios de muy diversos tipos.

¿Cómo pueden dividirse las pruebas e inventarios?

Fundamentalmente en dos tipos: estandarizados y *no* estandarizados.

Los primeros se caracterizan esencialmente porque poseen uniformidad en las instrucciones para administrarse y calificarse, así como ciclos específicos de desarrollo y reconfirmación. Los segundos, usualmente se generan mediante un proceso menos riguroso y su aplicación es limitada (como una prueba de un maestro para evaluar el aprendizaje de sus alumnos en una asignatura). Desde luego, los *no* estandarizados pueden irse robusteciendo (validándose en diferentes contextos) hasta convertirse en pruebas estandarizadas.

En teoría, un inventario estandarizado puede aplicarse a cualquier persona del tipo para el cual fue elaborado (Mertens, 2005) y el tiempo que se brinda para contestarlo en ocasiones es flexible y en otras no. Cabe destacar que solamente podemos comparar los resultados entre grupos a los que se les administró una prueba estandarizada (por ejemplo, en experimentos), cuando *no* existen diferencias entre los grupos que puedan generar sesgos en las respuestas (imaginemos que pretendemos comparar las respuestas a un *test* de razonamiento matemático entre niños de 10 años y niñas de seis años de edad, sería incorrecto). Por supuesto que, cuando contrastamos grupos (hombres frente a mujeres, entre distintos niveles socioeconómicos, jerarquías en la empresa, ocupaciones, entre otros casos) buscamos evaluar si hay o no diferencias, pero nos interesan las diferencias en la(s) variable(s) medida(s) o los resultados de la prueba. En cambio, *no* nos deben interesar las discrepancias entre grupos en las capacidades para el desempeño de ésta, porque supuestamente las capacidades son las mismas (salvo que el objetivo del estudio sea considerar disparidades en la ejecución de un inventario). El esquema que se pretende al comparar grupos podría esquematizarse como en la figura 7.3.

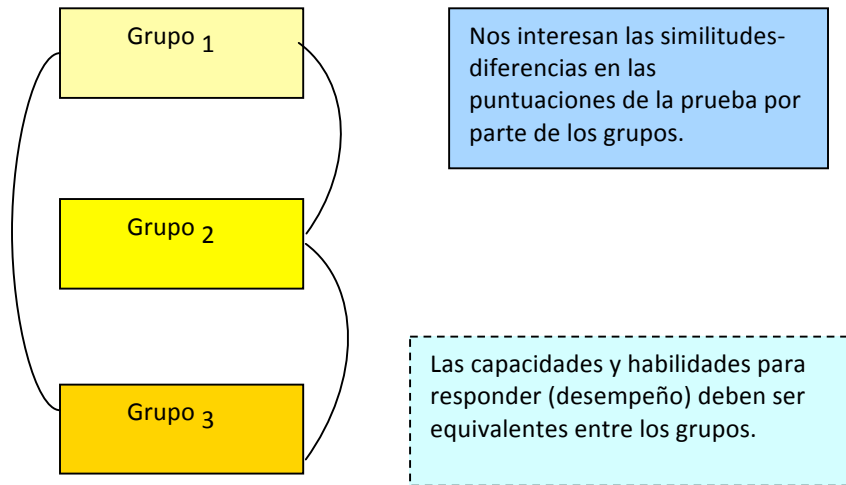


Figura 7.3 Interés al comparar grupos con la misma medición

Por ejemplo, si queremos comparar a varios pilotos y saber cuál es el mejor, entonces los autos deberán ser los mismos, la pista igual, el combustible, equipo humano, etc. Si probamos autos y pilotos, la pista será la misma, las condiciones ambientales también, etc. Imaginemos que queremos comparar a cuatro grupos en cuanto a su inteligencia, uno de españoles, otro de peruanos, un tercero de italianos y un cuarto de germanos, y la prueba está en alemán (¿los resultados reflejarán la inteligencia o el conocimiento del idioma? Más bien lo segundo).

¿Qué pruebas estandarizadas se encuentran disponibles?

En la actualidad existe una amplia diversidad de pruebas e inventarios desarrollados por diversos investigadores para medir un sinnúmero de variables. Estas pruebas tienen su propio procedimiento de aplicación, codificación e interpretación, además se encuentran disponibles en diversas fuentes secundarias y terciarias, así como en centros de investigación y difusión del conocimiento. Hay *pruebas para medir habilidades y aptitudes* (como habilidad verbal, razonamiento, memoria, inteligencia,

percepción, habilidad numérica), *la personalidad, los intereses, los valores, el desempeño escolar, la motivación, el aprendizaje, el clima laboral en una organización, la inteligencia emocional, etc.* También se puede disponer de *pruebas clínicas* para detectar conducta anormal, pruebas para seleccionar personal, pruebas para conocer las percepciones u opiniones de las personas respecto a diversos tópicos, pruebas para medir la autoestima y, en fin, otras muchas variables del comportamiento.

El reto en el empleo de estas pruebas es que la mayoría se han desarrollado en contextos muy diferentes al iberoamericano, por lo que en ocasiones su utilización resulta inadecuada, inválida y poco confiable. Cuando se utilice una prueba estandarizada como instrumento de medición, es conveniente que *se seleccione una desarrollada o adaptada por algún investigador para el mismo ámbito de nuestro estudio, que sea válida y confiable* (debemos tener información sobre este respecto).

En caso de que elijamos una prueba diseñada en otro contexto, será necesario adaptarla y efectuar pruebas piloto para calcular su validez y confiabilidad, así como ajustarla a las condiciones de nuestra investigación. Recordemos que cada vez que se administra un instrumento resulta indispensable demostrar que es válido y confiable para el ámbito en el cual se va a aplicar.

Como ya comentamos, Nuñez (2001) desarrolló una prueba estandarizada para medir el sentido de vida (bajo la concepción de Victor Frankl), una versión con 59 ítems, mediante el escalamiento tipo Likert, con el reporte de una confiabilidad alfa-Cronbach de 0.91, asimismo, con una solución de ocho factores, ubicados en cuatro cuadrantes de dos ejes: *homo faber* (éxito-fracaso) y *homo patiens* (plenitud-desesperación). Revisó, antes de crear su prueba, 15 instrumentos como el PIL, Logo Test y Song Test, Mile Test, Life Purpose Questionnaire Test, etc. Comenzó con un instrumento de 180 reactivos que depuró hasta llegar a sus 59 finales. Es un excelente ejemplo de una adaptación al medio latinoamericano de una prueba estandarizada sobre el sentido de vida. Los ítems definitivos se incluyen al terminar esta sección del capítulo.

En la actualidad encontramos gran variedad de estas pruebas en formato digital (CD y DVD) que hacen mucho más fácil y divertido contestarlas; un ejemplo es la prueba Emotional IQ Test (prueba de inteligencia emocional) de Meyer, Salovey y Caruso (1999). En internet hay diversos sitios donde se localizan estas pruebas: en su buscador teclee entre comillas la variable a medir y la palabra

“prueba o *test*”, y encontrará un sinnúmero de estos instrumentos para recolectar datos, pero haga una adaptación/contextualización de acuerdo con lo señalado en el capítulo 9 del libro impreso. También busque en los sitios recomendados en el capítulo 3 de este CD anexo.

Un tipo de pruebas estandarizadas bastante difundido lo constituyen las *pruebas proyectivas*, las cuales presentan estímulos a los participantes para que respondan a ellos; después se pueden analizar las respuestas tanto cuantitativa como cualitativamente, y se interpretan. Estas pruebas evalúan proyecciones de las personas, como por ejemplo la personalidad y sus trastornos (Rodríguez, 2007).

Dos pruebas proyectivas muy conocidas son el *test de Rorschach* (que presenta manchas de tinta en tarjetas o láminas blancas numeradas a los sujetos y éstos relatan sus asociaciones e interpretaciones en relación con ellas) (Weiner, 2009) y el *Test de Apercepción Temática* o TAT (que, con un esquema similar al de Rorschach, muestra a los participantes cuadros que evocan pasajes o historias, y ellos deben elaborar una interpretación, que es analizada por expertos).⁴ Por ejemplo, Pirkko (2008) examinó historias de 40 niños para evaluar sus fantasías y uso de la realidad y Blankenship et al. (2006) utilizaron el TAT para medir la necesidad de logro.

Dichas pruebas poseen un procedimiento estandarizado, aunque también se consideran instrumentos cualitativos para recolectar los datos, dada la profundidad de información que recogen de las personas en sus propias palabras. Cabe señalar que su aplicación regularmente es estandarizada; pero su interpretación no. El análisis es más bien de corte cualitativo.

Bastantes pruebas estandarizadas (como las proyectivas) requieren de un entrenamiento considerable y un conocimiento profundo de las variables por parte del investigador que habrá de administrarlas e interpretarlas. No pueden aplicarse con superficialidad e indiscriminadamente. La manera de aplicar, codificar, calificar e interpretar las pruebas estandarizadas es tan variada como los tipos existentes.

Algunos ejemplos de las pruebas, se incluyen en la tabla 7.6.

⁴ Para profundizar en la técnica y ver ejemplos de su aplicación se recomienda a: Langan-Fox y Grant (2006), Blankenship et al. (2006), Weertman *et al.* (2006), Sheldon et al. (2007) y Husain (2009).

Tabla 7. 6 Ejemplos de pruebas estandarizadas

Variable general	Variable específica	Instrumento/ referencias o ejemplos de aplicación
Inteligencia	Coeficiente de inteligencia o IQ en adultos	Escala de inteligencia de Wechsler (WAIS) (Lichtenberger y Kaufman, 2009; Hartman, 2009; Kaufman, 2009)
	Inteligencia sensorio-motriz de bebés (6 meses a 2 años)	Escala Argentina de Inteligencia Sensorio-motriz (Oiberman, Mansilla y Orellana, 2001; Caballero y Contini, 2008)
Personalidad	Personalidad en prácticamente todas sus dimensiones	Inventario Multifacético de la Personalidad (MMPI), buscar actualizaciones recientes lo más que sea posible (versión adultos y versión adolescentes) (Greene, 2009; Toomey, Kucharski y Duncan, 2009; Elliott, 2008; Dao, Prevatt y Horne, 2008)
Aptitudes mentales	Nivel general de aptitudes mentales en niños	Escala Stanford-Binet para niños (Wallace, 2009; Newton et al., 2008). En el caso de niños autistas se sugiere a Coolican, Bryson y Zwaigenbaum (2008)
	Aptitud cognitiva (nivel mental) para adolescentes y adultos	Escala Binet-Simon o Standfor-Binet (Wallace, 2009; Terman, 2008; Roid, 2003)
Actitud hacia el colegio	Actitud de adaptación en el colegio en estudiantes de educación media y superior	Inventario de adaptación al colegio WPS de Brown y Leigh (2009)
Creencias familiares	Creencias familiares de los padres	Inventario de las creencias familiares de los padres (Fischer y Corcoran, 2007)
	Creencias familiares de los adolescentes	Inventario de las creencias familiares de los adolescentes (Fischer y Corcoran, 2007)
Dolor	Dolor físico	Cuestionarios del dolor McGill [MPQ] (Ljunggren, Strand y Johnsen, 2007; Hyunjeong et al., 2008)
Satisfacción en el matrimonio	Satisfacción de la relación marital	Escala de felicidad marital [MHS] (Fischer y Corcoran, 2007)

Percepción de la vida	Satisfacción con la vida	Escala de satisfacción con la vida (SWLS) (Lopez y Snyder, 2003; Vassar, 2008; Gouveia et al., 2009)
Conciencia ecológica	Conciencia ecológica	Escala de nueva conciencia ecológica de Ellis y Thompson (Robinson, Shaver y Wrightsman, 1999)
Calidad en la práctica de la enfermera	Intervenciones para asistir al paciente en estado de estrés	Inventario de intervenciones para asistir al paciente en situaciones estresantes de Edelman y Mandle (Stanhope y Knollmueller, 2001)
Dependencia	Dependencia interpersonal	Inventario de dependencia interpersonal de Hirschfield (IDI) (Tzeng, 1993; Shahar, 2008). Así como también la subescala del Inventario del Estilo Personal de Robins (PSI) (Shahar, 2008)
Lealtad del cliente	Lealtad hacia la marca	Escala de Wunderman de experiencia con la marca (Morel, 2003)
Jubilación	Actitud hacia la jubilación	Escalas para medir las actitudes hacia la jubilación usando diferencial semántico y escalamiento de Thurstone (Rodríguez, 2006)
Matrimonio	Actitud hacia el matrimonio	Actitud e integración hacia el matrimonio (Riggio y Weiser, 2008)
Inteligencia emocional	Inteligencia emocional y personalidad	20 cuestionarios de personalidad (rasgos) (Carter, 2009), así como pruebas de inteligencia emocional (EI) (Stough, Saklofske y Parker, 2009). En España se adaptó el Meta Mood Scale-24 para medir la IE (Pérez y Castejón, 2007)
Mascotas	Actitud hacia las mascotas	Escala de actitud hacia las mascotas (Morovati et al., 2008)
Comportamiento de compra	Comportamiento de compra compulsiva	Compulsive Buying Scale (Escala de compra compulsiva) (Koran et al., 2007)

Hay miles de pruebas e inventarios en las fuentes secundarias y bases de datos que se mencionan en los apéndices 1 y 2 de este CD, así como páginas en internet como:

- Tests and Measures in the Social Sciences: Tests available in compilation volumes (pruebas de diversas clases): <http://libraries.uta.edu/helen/test&meas/testmainframe.htm>
- PsycInfo: <http://www.apa.org/psycinfo/>
- TestLink (ETS Test Collection Advanced Search o Tests on Microfiche) y The Mental Measurements Yearbook (también en versión manual de base datos). Se ingresa por medio de las páginas de las universidades (debe su institución estar suscrita).
- ERIC (mencionada ampliamente): <http://www.eric.ed.gov/>
- Psych Tests, BBC Surveys and Psychology Tests o Queendom.com.

Adicionalmente se recomiendan: Kubiszyn y Borich (2009), Barrett (2009), Anastasi y Urbina (2009), Groth-Marnat (2009), Abell, Springer y Kamata (2009), Thorndike y Thorndike-Christ (2009), Kaplan y Saccuzzo (2008), Mertens (2005) y Mertens y McLaughlen (2004) (particularmente pruebas para personas con capacidades diferentes), Kamphaus y Reynolds (2003), Thorndike y Hagen (1989), Cronbach (1984), Norton (1980), Woelfel y Danes (1980) y, desde luego, las obras clásicas de Torgerson (1985) y Nunnally (1970). Se trata de autores especializados sobre medición y manejo de pruebas estandarizadas.

Además, existen métodos propios de las diferentes ciencias sociales, como el análisis de redes, para evaluar la forma en que se manifiesta la comunicación en un sistema social (quién se comunica con quién, quiénes distorsionan la información, cómo fluye la comunicación, quiénes son los líderes comunicativos, etc.), sistemas de medición fisiológica, escalas multidimensionales que miden a los sujetos en varias dimensiones, como los sistemas Galileo y ALSCAL de Woelfel y Fink (1980), así como Borg y Groenen, 2005), para medir la distancia psicológica entre conceptos en individuos (como por ejemplo: “patria”, “madre”, “presidente”, “nación” y cualquier otro), en los que se consideran dimensiones cognitivas y emocionales. En fin, otros métodos más que escapan del alcance de este libro.

Para el análisis de redes, se sugiere a Rogers y Kincaid (1981), Wellman y Berkowitz (1997), Wellman (1999), Knoke y Yang (2007), y Hassanien, Abraham y Snasel (2009). Un ejemplo de esta clase de investigación es la de Crossley (2008), quien analizó la red social de los “actores clave” involucrados en el “círculo central” del movimiento *punk* en el Reino Unido.

Ejemplo de ítems de prueba estandarizada (Sentido de vida Celaya)⁵

Los ítems o afirmaciones son calificados a través de una escala Likert que va desde “totalmente de acuerdo” hasta “totalmente en desacuerdo”.

1. En mi familia nos escuchamos, nos demostramos amor y ocasionalmente analizamos los desacuerdos.
2. Mi religión ha fomentado el crecimiento de mi capacidad de amar y desarrollarme.
3. Dedico parte de mi vida a realizar actividades desinteresadamente.
4. Estoy convencido de que estoy bien; me acepto y me amo.
5. En la búsqueda de la felicidad intento lograr satisfactores materiales.
6. Honestamente tengo conciencia de que amo las cosas y utilizo a las personas.
7. La afectividad es muy importante en mi vida.
8. Mi alto nivel de energía lo empleo para triunfar.
9. Mi religión me ha proporcionado razones para vivir con rectitud, para estar bien con Dios.
10. Acepto que me traten indignamente, con tal de tener una seguridad económica.
11. Promuevo que las personas que dependan de mí tengan aprendizajes significativos.
12. Me gusta todo lo que es fácil y que no requiere pensar mucho.
13. Asisto a congresos, conferencias, tomo cursos y diplomados.
14. Me gusta comunicarme conmigo y sentir que mi cabeza sabe lo que tiene mi corazón.

⁵ Núñez (2001). Este estudio puede verse a manera de protocolo o proyecto de tesis en este CD: Material complementario→Investigación cuantitativa→Ejemplo 5: “Exploración del sentido de vida a través del diseño de un instrumento cuantitativo”.

15. Le agrado a la mayoría de las personas.
16. Si me dieran la oportunidad, podría hacer algo que realmente ayudara a otros.
17. Repentinamente mi estado de ánimo cae en la desesperanza.
18. Valoro todo lo que favorece mis posibilidades de estar bien.
19. Puedo vislumbrar la posibilidad de cambiar y de crecer en mi persona.
20. Me gusta comprometerme en obras de tipo social.
21. Contagio mi optimismo.
22. Mis problemas personales influyen en mi trabajo.
23. En mi proyecto de vida establezco metas relacionadas con un aprendizaje significativo.
24. En mi familia muchas veces hemos rezado juntos.
25. Siento que mi cuerpo pierde fortaleza interna.
26. He buscado activa y libremente a Dios.
27. A mis familiares les disgusta comprometerse en obras de tipo social.
28. En mi familia mis padres acostumbraban devaluarse unos a otros.
29. Me siento insatisfecho con mi trabajo.
30. Mis sentimientos me impiden ver con claridad la realidad.
31. Tengo la sensación de que algo le falta a mi vida.
32. A la persona que amo la acepto incondicionalmente.
33. Mi meta al educar no sólo es informar, sino afinar la conciencia.
34. En estos momentos me doy cuenta de que no puedo ocuparme de nada porque no tengo voluntad de hacerlo.
35. Promuevo una educación que tenga una relación más estrecha con la vida.
36. Invierto parte de mis ingresos en adquirir revistas científicas, buenos libros, videos, documentales, enciclopedias.
37. Con frecuencia siento que algo malo me va a ocurrir.
38. Mis opiniones políticas no están bien definidas.
39. Renuncio a mi libertad para obtener satisfactores que me convengan.
40. Soy sensible al sufrimiento de los demás.
41. Trato de estar con personas divertidas para que me ayuden a salir de mi aburrimiento.
42. En mi familia se carece de sentimientos de lealtad y de pertenencia.

43. Me siento lleno de vida y vitalidad.
44. Disfruto la lluvia y no me importa mojarme.
45. Me parece que sufro más que los demás cuando rompo una relación.
46. Prefiero mantenerme a cierta distancia de la gente y quedarme en mi soledad.
47. Me considero una persona optimista.
48. La vida es como una pizza llena de ingredientes apetecible y deliciosa.
49. Me siento capaz para trabajar eficientemente.
50. Soy instrumento de las circunstancias y el destino.
51. No pasa un día en que no dedique unos minutos para contemplar la belleza de la naturaleza.
52. En mi trabajo confío en mis dotes y en mi capacidad de adaptarme para hacer frente a los desafíos que se presentan.
53. Admiro a personas como Martin Luther King, Mandela y la madre Teresa de Calcuta.
54. Diariamente trato en mi trabajo de dar lo mejor de mí mismo.
55. Tengo necesidad de protagonismo, de ser notado.
56. Me gusta echarme boca arriba en el césped y sacar figuras de las nubes.
57. Hay en mi interior una gran riqueza y sensibilidad.
58. Puedo permitirme vivir sentimientos delicados o tiernos.
59. Me siento capaz de hacer bien mi trabajo.

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 comunicaciones en sus múltiples formas: verbales, no verbales, escritas y audiovisuales.

Análisis de contenido

Técnica para estudiar la comunicación de una manera objetiva, sistemática y que cuantifica los contenidos en categorías.

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 y 2003) extiende la definición del *análisis de contenido* a un método de investigación para hacer inferencias válidas y confiables de datos sobre contenidos con respecto a su contexto. Weber (1990) y Krippendorff y Bock (2008) señalan que tales inferencias implican a la fuente de los mensajes, a los mensajes en sí mismos y a sus receptores o destinatarios.

Resulta una técnica muy útil para analizar los procesos de comunicación y sus características en muy diversos contextos (Neuendorf, 2001). El análisis de contenido puede aplicarse virtualmente a cualquier forma de comunicación (páginas web y blogs, 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, evaluando sus escritos; conocer las actitudes de un grupo de personas mediante la valoración de sus expresiones verbales; indagar sobre las preocupaciones de un pintor o un músico al estudiar su material; compenetrarse con los valores de una cultura; o averiguar las intenciones de un publicista o propagandista.

Usos del análisis de contenido

Berelson (1971) y Weber (1990) señalan 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 enseñanza, el aprendizaje de conceptos).*
14. *Anticipar respuestas a comunicaciones.*
15. *Reflejar actitudes, valores y creencias de personas, grupos o comunidades.*
16. *Cerrar preguntas abiertas.*

Asimismo, puede resultar útil para examinar redes sociales (quiénes se comunican con quiénes —perfiles—, para qué —propósitos—, con qué frecuencia, etc.). Por ejemplo, podría estudiarse una red de jóvenes inscritos en Facebook, Hi5 o MySpace, mediante el análisis de sus páginas (mensajes escritos e imágenes).

¿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 en 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* (Krippendorff y Bock, 2008).

Universo

El **universo** podría ser la obra completa de Paulo Coelho o Stephenie Meyer; 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* y *Diario 16*); todos los capítulos de tres telenovelas; las canciones completas de un grupo o solista (U2, Joan Manuel Serrat, Shakira, Coldplay, La Oreja de Van Gogh, Motel, 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. 221), 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 señalan Kerlinger y Lee (2002) y Franzosi (2009), puede haber unidades más pequeñas, como letras, fonemas, símbolos. Así, se puede medir cuántas 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”, “mí” y otros términos que indiquen referencia al “yo” del escritor. Así, se analizaría qué 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.
3. *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

(Weber, 1990). Ejemplos de ítems son un libro, un editorial, un programa de radio o televisión, un discurso, una ley, 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.

4. *El personaje*. Se trata de un individuo. Por ejemplo, un paciente, un héroe o heroína de una serie televisiva, un líder histórico, etc. Aquí lo que se analiza es el personaje.

5. *Medidas de espacio-tiempo*. Son unidades físicas, como el centímetro-columna (por ejemplo, en la prensa), la línea (en escrito), el minuto (en una conversación o en radio), el mensaje (enviado por celulares o móviles), el periodo de 20 minutos (en una interacción), el cuadro (en televisión), cada vez que se haga una pausa (en un discurso), o una página (blogs). Sánchez Aranda (2005) lo 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 7. 4.

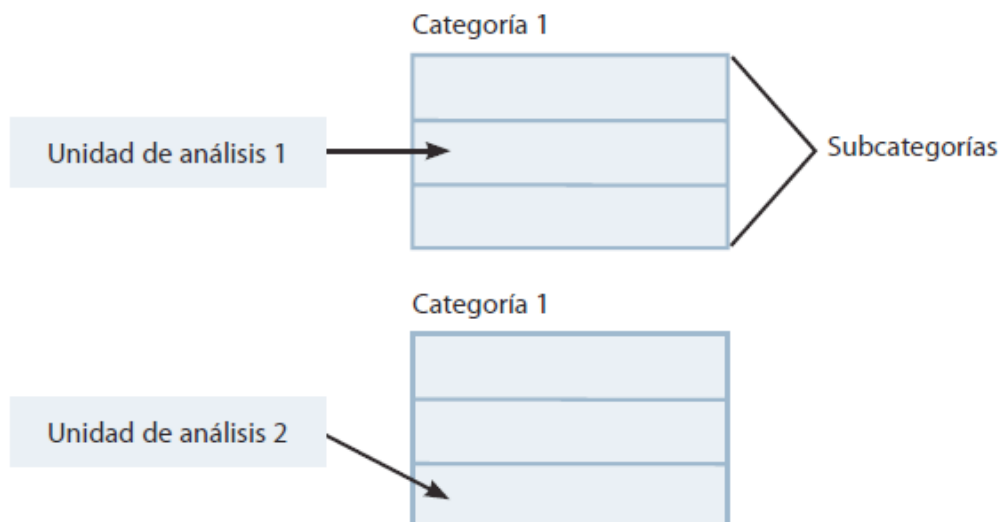


Figura 7.4 Agrupación de unidades de análisis en categorías

Al cuestionarnos sobre qué unidad debe seleccionarse, esto depende de los objetivos y las preguntas de investigación. Sin embargo, Berelson (1971), así como Krippendorff y Bock (2008) sugieren lo siguiente:

- a) En un solo estudio es posible utilizar *más de una unidad de análisis*.
- b) Los cálculos de palabras y las unidades amplias, como el ítem y las medidas de espacio-tiempo, son más 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.
- 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. Weber (1990) y Sánchez Aranda (2005) las definen 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 catalogaría como bueno, neutral o malo. En ambos casos, la unidad de análisis se categoriza. Veamos esquemáticamente la categorización en la figura 7. 5.

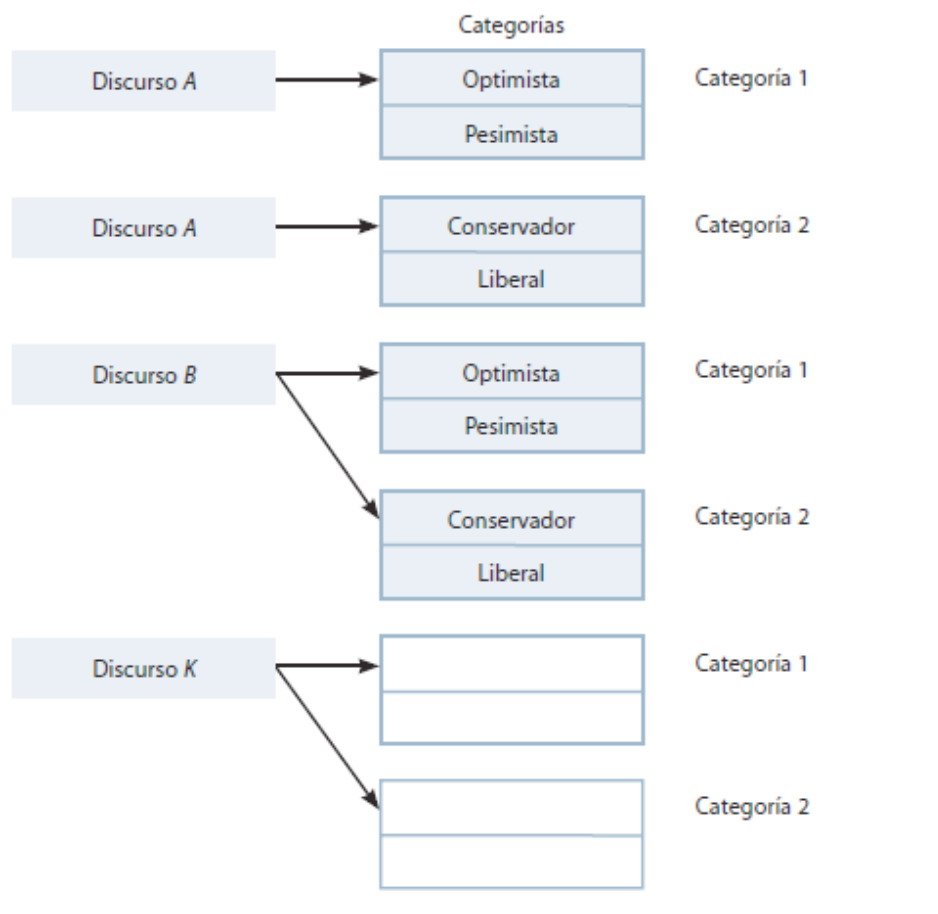


Figura 7.5 Ejemplos de unidades de análisis y su clasificación en categorías

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 7. 6.

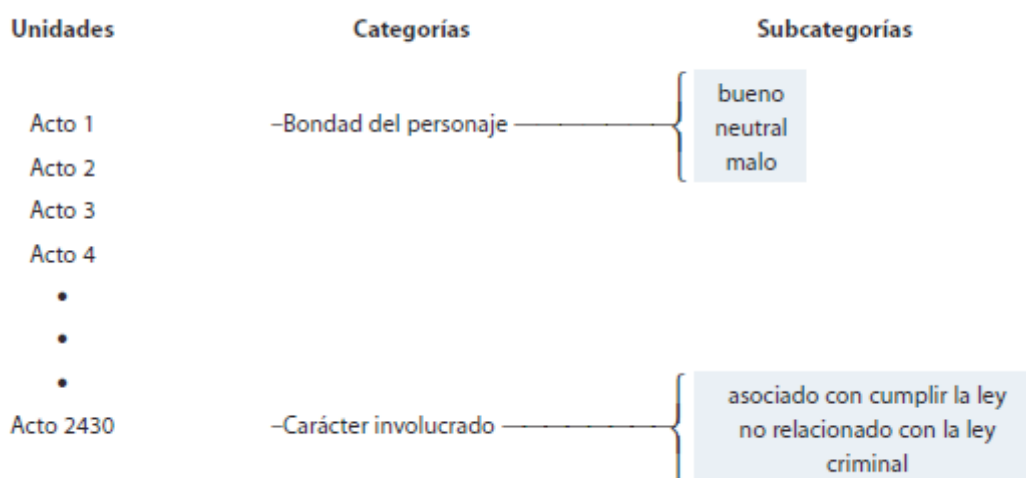


Figura 7.6 Ejemplo con dos categorías y seis subcategorías

La selección de categorías también depende del planteamiento del problema.

Tipos de categorías

Krippendorff (2003) señala cinco tipos básicos de categorías:

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

EJEMPLO

Analizar las problemáticas que más preocupan a los ciudadanos detectadas en los mensajes enviados por correo electrónico al alcalde o presidente municipal durante el último mes.

Categoría:	Inseguridad
Subcategorías:	Asaltos/robos al transitar o caminar en las calles o lugares públicos abiertos, como parques
	Asaltos/robos al transitar en automóviles
	Asaltos/robos al viajar en transporte público
	Asaltos/robos a los hogares
	Asaltos/robos en lugares públicos cerrados, como restaurantes, cines, etcétera
	Secuestros
	Violaciones
	Etcétera
Categoría:	Servicios públicos proporcionados por el municipio
Subcategorías:	El servicio de recolección de la basura no ocurre todos los días
	El servicio de recolección de la basura no se realiza más o menos a la misma hora todos los días, es irregular
	Cortes frecuentes en el suministro de agua
	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 del presidente de Estados Unidos.

Categoría:	Tono en el tratamiento respecto al presidente de Estados Unidos
Subcategorías:	Favorablemente
	Neutral
	Desfavorablemente

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

EJEMPLO

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

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

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

EJEMPLO

Analizar a quiénes encauza más sus críticas el rector o director general de la universidad en los discursos del último año dirigidos a toda la comunidad estudiantil.

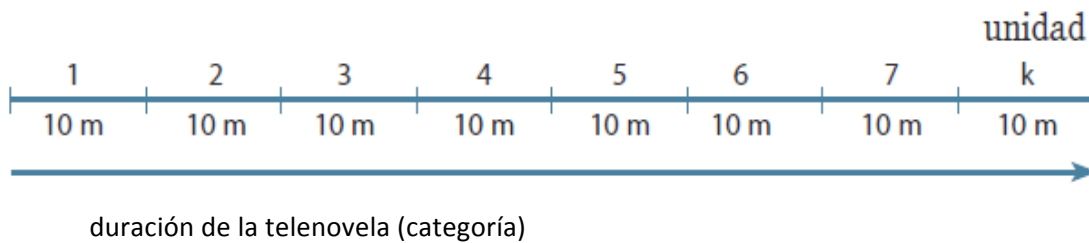
Categoría:	Receptores a quienes se dirigen las críticas del rector o director
Subcategorías:	Profesores de tiempo completo
	Profesores que no están de tiempo completo (de asignaturas, por honorarios o tiempo parcial)
	Alumnos en general
	Alumnos de últimos semestres
	Alumnos de semestres intermedios
	Alumnos de primeros semestres
	Directivos
	Empleados administrativos
	Padres de familia
	A toda la comunidad

5. *Físicas*. Son categorías para ubicar la posición y duración o extensión de una unidad de análisis. De *posición* son, por ejemplo, la sección, página y cuadrante (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 o escenas en una película, etcétera.

EJEMPLO

Imaginemos que pretendemos investigar el grado en que una determinada telenovela presenta consumo de alcohol entre sus personajes y las situaciones y contextos en los cuales ocurre, así como su finalidad (reducir estrés, inducir al romance, socializar, etcétera).

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



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 7.7 con un par de ejemplos simples.

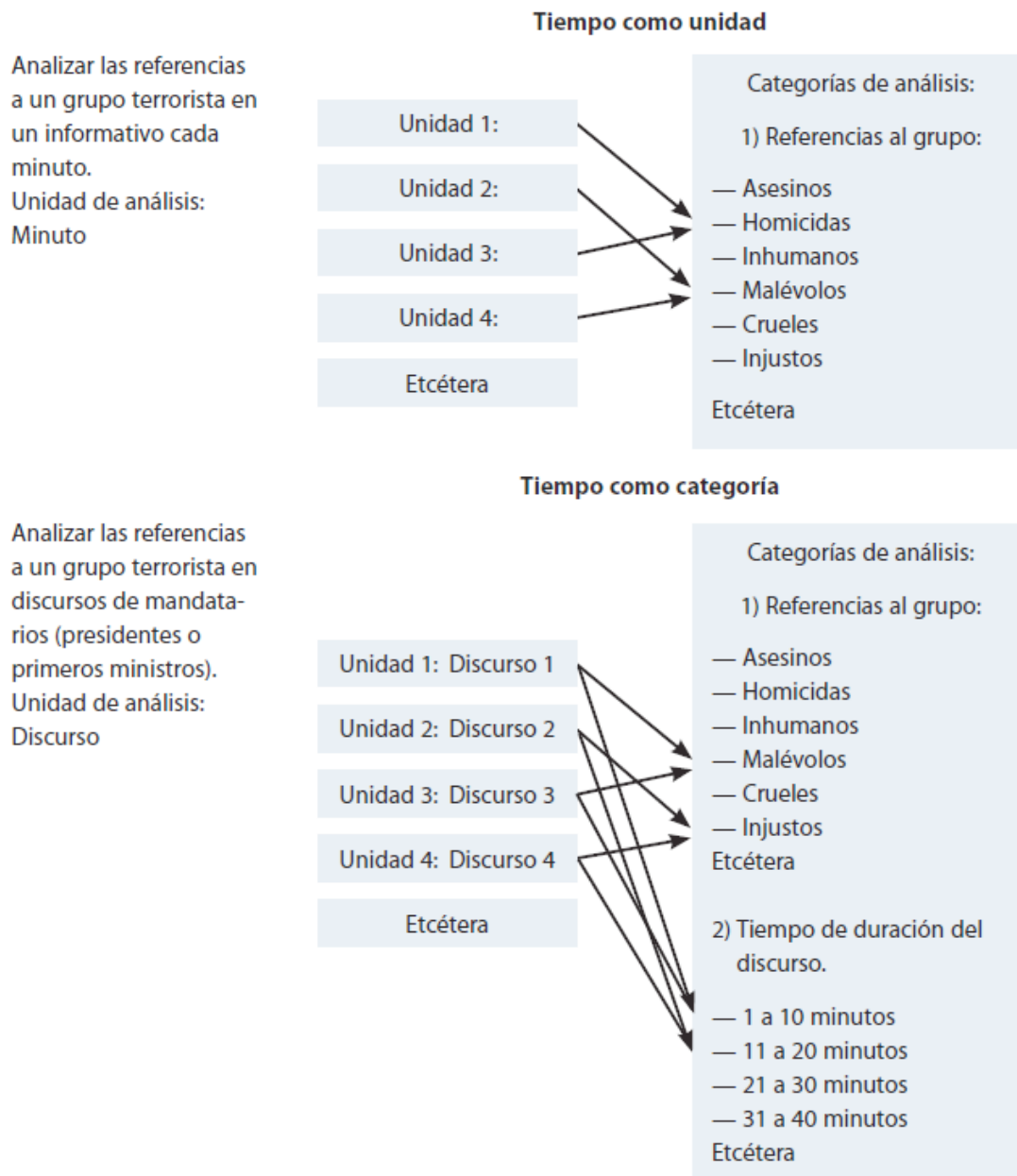


Figura 7.7 Ejemplos del tiempo como unidad de análisis y como categoría

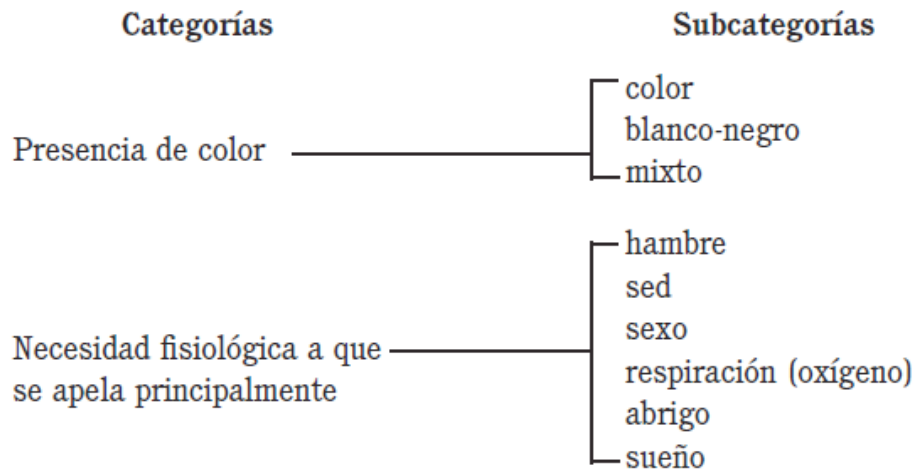
Un sexto tipo de categorías es la de “intensidad”, que indica más allá de la dirección, grados de *fuerza* de la categoría. Por ejemplo, diversos gabinetes, despachos o departamentos de comunicación social de organismos públicos (como sería el caso de ministerios o secretarías), examinan diariamente las críticas que hacen los editorialistas

de los diarios o periódicos respecto al titular de la dependencia (ministro o secretario), y determinan si la crítica es: “muy severa”, “severa” y “no severa”.

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 de una serie televisiva no puede ser “bueno” y “malo” a la vez.
3. 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, etcétera).

Un comercial podría utilizar más de una apelación (salud y belleza, por ejemplo).

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

EJEMPLO

De un análisis de contenido de series televisivas

Para ejemplificar el análisis de contenido y específicamente la generación de categorías, se acudirá a un estudio de Greenberg et al. (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 7. 8.

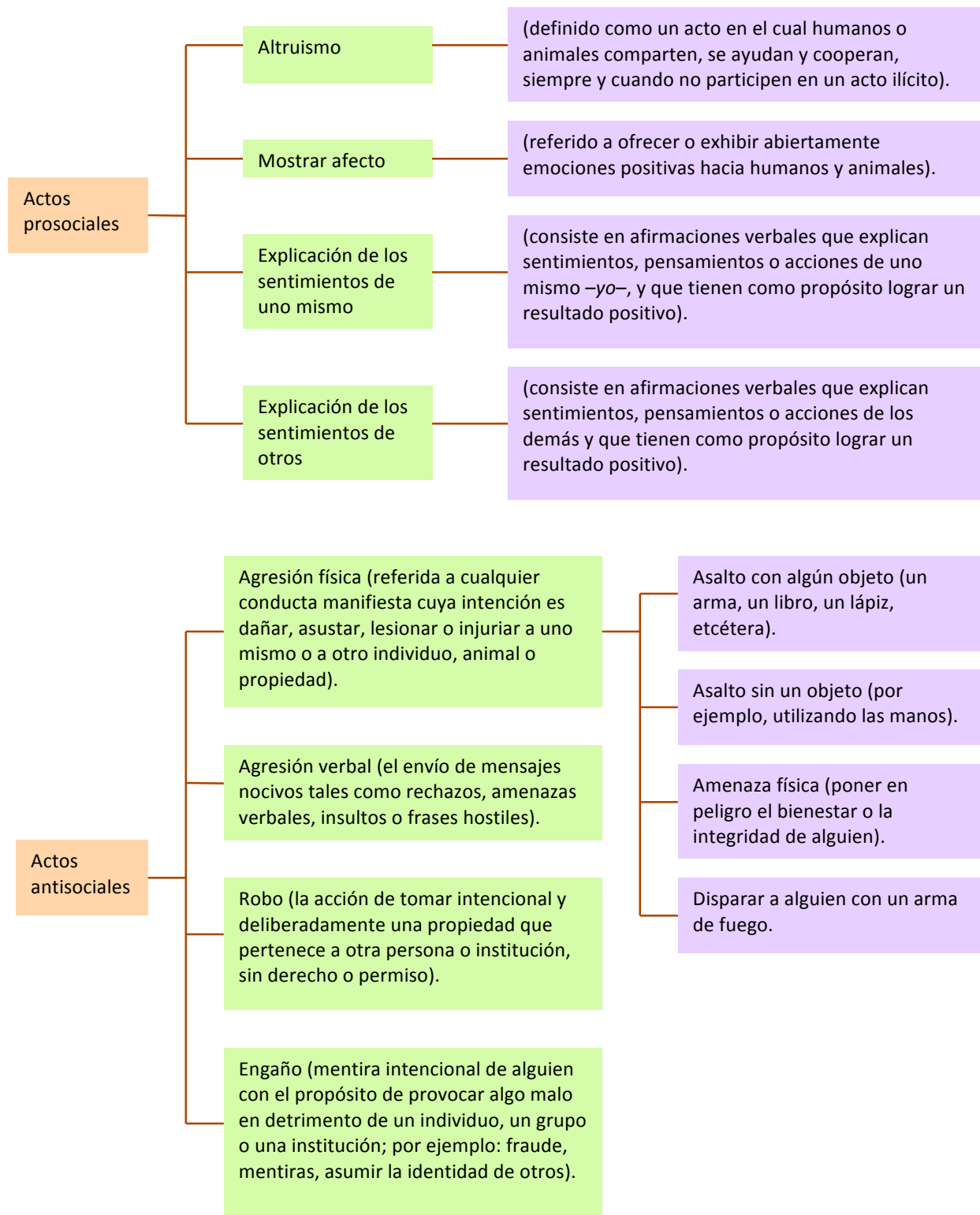


Figura 7.8 Ejemplo de sistema de categorías

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 7. 7.

Tabla 7. 7 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*	
	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
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	1 629	(55.0)	1 099	(47.6)	1 464	(54.0)
III. Robo	61	(2.1)	72	(3.1)	44	(1.6)
IV. Engaño	<u>283</u>	(9.5)	<u>251</u>	(10.9)	<u>319</u>	(11.8)
Total	2 964		2 308	(100.0)	2 713	(100.0)
Horas analizadas	(68.5)		(58)		(63)	
*Año 1 incluyó de octubre de 1975 hasta que se grabó un episodio de cada una de las series existentes (1976), año 2 igual, pero en 1976-1977, y año 3 igual, pero de 1977-1978. <i>f</i> = número de casos o unidades. % = porcentajes.						

A continuación mostramos un ejemplo de análisis de contenido, a manera de ejercicio, que se hace cotidianamente en la mayoría de los países para clasificar las películas y regular el ingreso (por el factor edad) a las salas cinematográficas para verlas.

EJEMPLO

Ejercicio

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 sólo son apropiadas para adolescentes y adultos, o únicamente son aptas para adultos (mayores de 18 años). Por supuesto, la unidad de análisis es el filme o cinta, pero: ¿qué categorías podrían utilizarse para clasificar las unidades?

Algunas de las categorías para clasificar 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 irreversibles 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:

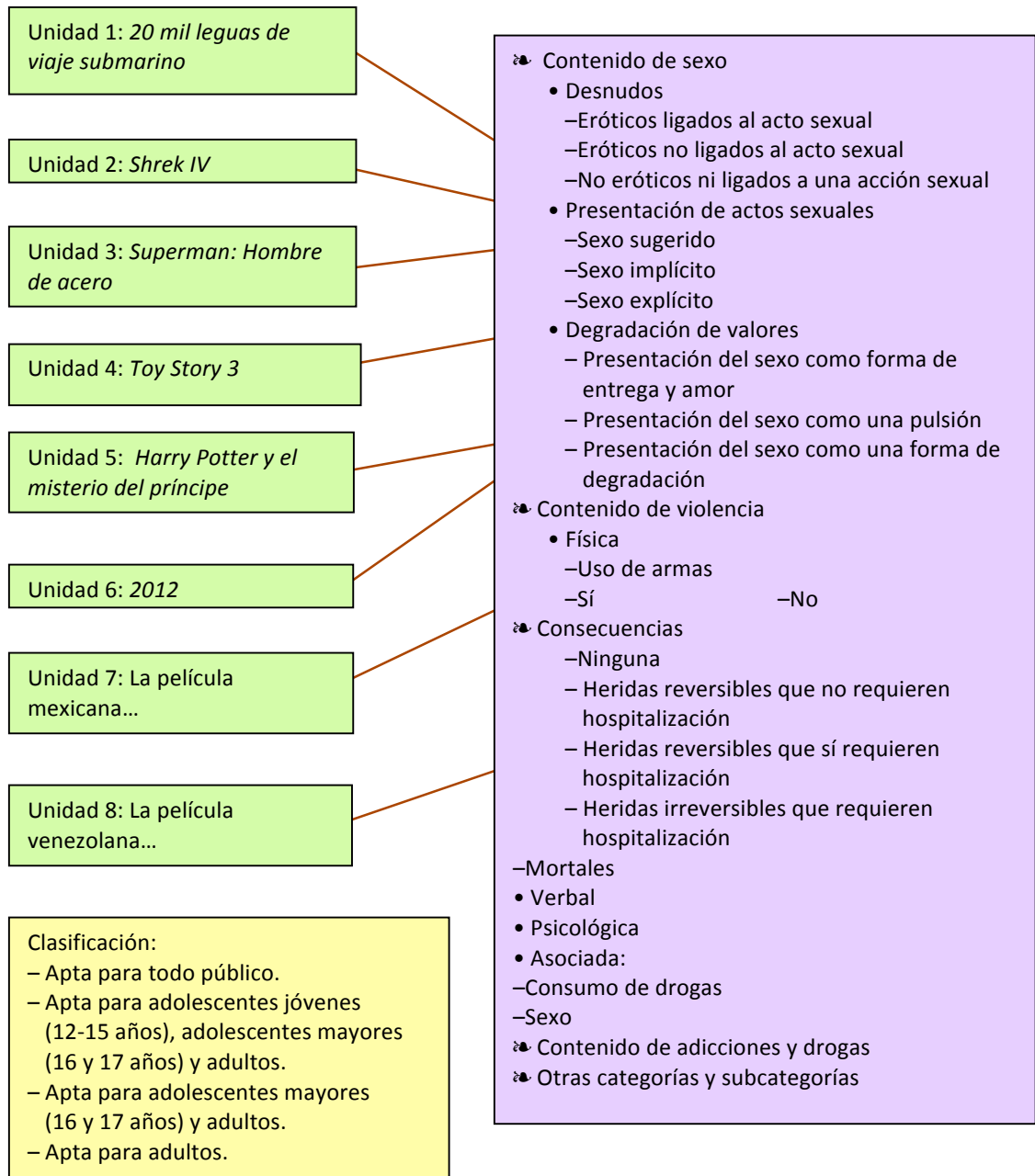


Figura 7.9 Ejemplo de un proceso de categorización de películas

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.

¿Cuáles son los pasos para llevar a cabo el análisis de contenido?

Ya hemos mencionado tres:

1. *Definir con precisión el universo y extraer una muestra representativa.*
2. *Establecer y definir las unidades de análisis.*
3. *Establecer y definir las categorías y subcategorías que representan las variables de la investigación.*

Los demás pasos son:

4. *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 de licenciatura, como mínimo).
5. *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 7. 10 se muestran los elementos de una hoja o registro de codificación, que puede estar impreso en papel, pantalla de PC o palm. Incluso puede ser una hoja en Excel.

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		MATERIAL A ANALIZAR:		CAPÍTULO DE LA SERIE "LOST" (PERDIDOS)
FECHA: 29-XI-08		DURACIÓN: 60 MINUTOS		
CATEGORÍAS		FRECUENCIAS		TOTALES
ACTOS PROSOCIALES (1)	ALTRUISMO (1.1.0)			2
	MOSTRAR AFECTO (1.2.0)			5
	EXPLICACIÓN DE LOS PROPIOS SENTIMIENTOS (1.3.0)			3
	EXPLICACIÓN DE LOS SENTIMIENTO DE OTROS (1.4.0)			0
ACTOS ANTISOCIALES (2)	AGRESIÓN FÍSICA 2.1	Asalto con objeto (2.1.1)		0
		Asalto sin objeto (2.1.2)		0
		Amenaza física (2.1.3)		0
		Disparar (2.1.4)		0
		Otros (2.1.5)		0
	AGRESIÓN VERBAL (2.2.0)			3
	ROBO (2.3.0)			0
	ENGAÑO (2.4.0)			1
			TOTAL	14

Categorías y subcategorías

Suma de frecuencias en cada categoría o subcategoría

Total de frecuencias o unidades

Al reverso pueden solicitarse comentarios del codificador, o bien si los análisis se efectúan en PC se hacen en comentarios (Word) o en la parte adecuada para ello de Excel.

Figura 7.10 Elementos de una hoja o registro de codificación

Las *hojas o registros 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,

X	

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 7.10, 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 sólo 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	Codificador B	Codificador C	Codificador D
0.89	0.93	0.92	0.67

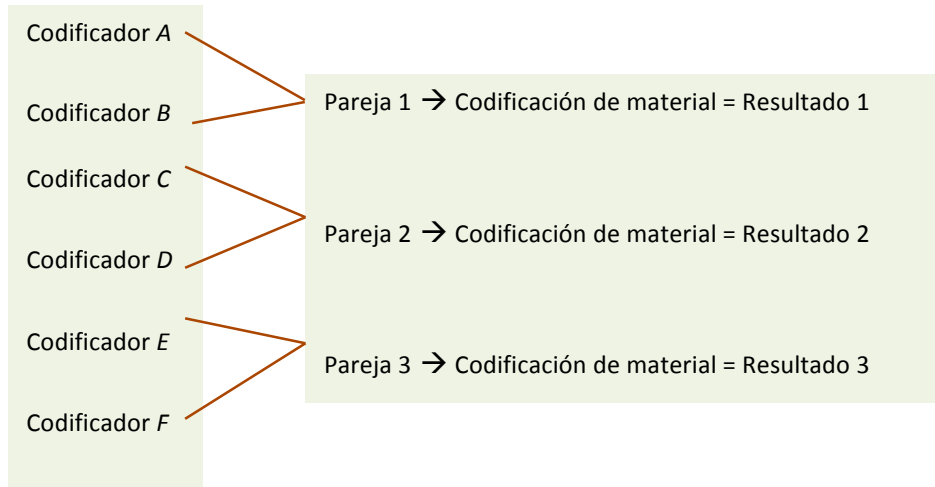
El codificador D tiene baja confiabilidad.

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

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

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

$$C_{1y2} = \frac{17}{18} = 0.94$$

$$C_{1y3} = \frac{16}{18} = 0.89$$

$$C_{2y3} = \frac{16}{18} = 0.89$$

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 16). 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>	
<i>A</i>	25	se toma el máximo	21
<i>B</i>	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:

$$C_T = \frac{C_{1y2} + C_{1y3} + C_{2y3}}{3}$$

$$C_T = \frac{0.94 + 0.89 + 0.89}{3} = 0.906$$

No es conveniente tolerar una *confiabilidad* menor de 0.85 (ni total, ni entre dos parejas) y de ser posible debe superar el 0.89. Al igual que con otros instrumentos de medición, la confiabilidad oscila entre 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 50%) y al completar ésta.

8. *Llevar a acabo 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).
9. *Vaciar los datos de las hojas o registros de codificación y obtener totales de cada categoría.* Este vaciado puede ser de las hojas o registros a la matriz de SPSS, Minitab o programa equivalente.
10. *Realizar los análisis estadísticos apropiados.* Para esto último, debe revisarse el siguiente capítulo.

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). Puede aplicarse creativamente a una enorme variedad de investigaciones. Por ejemplo, DeCicco (2007) estudió el contenido de los sueños de un grupo de mujeres y demostró su relación significativa con el descubrimiento de sí mismas. Tian (2009) examinó las apelaciones publicitarias utilizadas y los valores culturales reflejados en los anuncios en los periódicos y la televisión de China. Smith (2008) lideró un libro sobre múltiples aplicaciones del análisis de contenido para efectuar inferencias sobre la personalidad y la motivación de las personas, así como respecto a características y experiencias de grupos.

Finalmente, hemos de señalar que con el desarrollo de las nuevas tecnologías de comunicación, el análisis de contenido se mantiene vigente y en expansión.⁶

OBSERVACIÓN

¿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), así como Polster y Collins (2007) mencionan 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

⁶ Para revisar técnicas específicas aplicadas a los medios de comunicación masiva actuales, recomendamos a Divakaran (2009).

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, los riesgos en el trabajo, etcétera.

Observación cuantitativa

Registro sistemático, válido y confiable de comportamiento o conducta manifiesta. Se recolecta información sobre la conducta más que de percepciones.

La **observación cuantitativa**, como método para recolectar datos, es similar al análisis de contenido. Por ello, 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 (Bentzen, 2008; 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 un sistema y reglas que se aplican invariablemente, y deben minimizar su efecto sobre los registros e interacciones con los participantes observados. Esto es particularmente crítico porque los seres humanos somos selectivos en nuestras observaciones e inducimos cierto grado de variación (Ratcliff y Lassiter, 2007).

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 precisar lo que concebimos como “recurso escolar”. Un universo de aspectos a observar 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.

2. *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 para 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.⁷ 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). Asimismo, se debe 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,

⁷ 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).

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.

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 7.8 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: en el caso de cada cliente, desde que ingresa a la tienda y hasta que sale de ésta: ¿qué prendas o accesorios compró?, ¿de qué colores, texturas, ¿cuál es su perfil?, 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 puntualizadas para el análisis de contenido. Y la observación también consiste en asignar unidades a categorías y subcategorías de observación. En ésta, como en las demás técnicas 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 7.9. 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.
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 inquirirle 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 satisfacción con esa solución.

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	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 inquirirle si está satisfecho. Tiempo que el cliente espera en fila (cola).
--	--------------------------	---

EJEMPLO

De unidades y categorías

En el caso del experimento citado en el capítulo 7 (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 segundos. 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 participante 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 sub-

categorías 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 participante.
3. Conducta visual del sujeto: la cual 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 dicótomas (respuesta de evitación) son los silencios no naturales en el discurso, cuando expresamente el individuo se queda en silencio. La conducta verbal se mide a través del diálogo que sostenga la persona 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 individuo 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 con el planteamiento del problema e hipótesis.

Las subcategorías pueden ser escalas de actitudes

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

EJEMPLO

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 : — : — : — : — : — : — : Improductividad (5) (4) (3) (2) (1) (0)		
Actitud de servicio : — : — : — : — : — : — : Actitud de no servicio (5) (4) (3) (2) (1) (0)		
Cumplimiento-incumplimiento Etcétera		

6. *Diseñar la manera de registrar las observaciones* (mediante un formulario, hoja de codificación o guía de observación).⁸ 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 ocasiones, los instrumentos son aparatos que registran las observaciones (pistola láser de lectura de velocidad, scanner, etcétera). En el caso de Naves y Poplawsky (1984), el formato de codificación es el que se muestra en la figura 7.11, asimismo, se agrega otro ejemplo en la figura 7.12.

⁸ En papel o como documento de PC, puede ser una hoja de Excel.

Hora				Nombre del participante										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 7.11 Primer ejemplo de un formato de registro o codificación

El ejemplo de la figura 7.12 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).

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...), y “1” si son otras. Asimismo, en la fila “N” se indica el número de estudiantes involucrados en el episodio.

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 7.12 Formato de registro de conductas inapropiadas

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.

Anastas (2005) y también Polster y Collins (2007) plantean 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 presencia 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 (2001) 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, deben conocer las variables, categorías y subcategorías. Las habilidades más importantes de un buen observador son:
- Selectividad para registrar lo que observa, de acuerdo con reglas.
 - Uso de todos los sentidos, particularmente ante lo que no se esperaba encontrar.
 - Capacidad para seguir el sistema de registro que involucra 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, 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 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) y Grinnell, Williams y Unrau (2009) consideran como una confiabilidad interobservadores mínima aceptable 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).

11. Llevar a cabo la codificación por observación.
12. Vaciar los datos de las hojas o formatos de codificación y obtener totales para cada categoría.
13. Realizar los análisis apropiados.

CODIFICACIÓN EN EL ANÁLISIS DE CONTENIDO Y LA OBSERVACIÓN

Una vez recolectados los datos del análisis de contenido y/o la observación, también éstos deben codificarse, es decir, asignar códigos (generalmente numéricos) a las categorías y/o subcategorías. Por ejemplo, en el caso del estudio de Naves y Poplawsky (1984), la codificación es la que se muestra en la tabla 7.10 (que corresponde al primer ejemplo de un formato de registro o codificación, figura 7. 11).

Tabla 7.10 Ejemplo de codificación en el caso del estudio de Naves y Poplawsky

Variable	Categorías	Subcategorías	Códigos	Columnas
– Tratamiento experimental – Conducta	– Grupo cultural		1	1
	– Grupo sociopsicológico		2	
	– 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
	– MRM		2	
	– APY		3	

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.

Y la matriz de SPSS o equivalente sería la de la figura 7.13 (desde luego, solamente se incluyen 10 casos a manera de ejemplo).

	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 7.13 Matriz o vista de los datos para la codificación de la tabla 7.10 (Naves y Poplawsky)

RECOPILACIÓN Y ANÁLISIS DE DATOS SECUNDARIOS

La recopilación y análisis de datos secundarios se fundamenta en información disponible recolectada por otras personas distintas al investigador. Esta técnica para recabar los datos puede subdividirse en: *a)* datos estadísticos provenientes de fuentes oficiales y *b)* datos para metaanálisis. En ambos casos se acude a bases de información que pueden encontrarse en una o varias fuentes y que son útiles para los propósitos del estudio. La diferencia entre estos dos tipos estriba en que los primeros recogen estadísticas (tasas de defunción o mortalidad y nacimiento, número de hectáreas de riego, cifras de egreso en los distintos niveles del sistema escolar, número de camas de los hospitales, número de monumentos históricos, producto interno bruto, número de instalaciones deportivas y clases, así como número de cada una y tamaño en metros cuadrados, extensión de la red de alcantarillado, índice de inflación, monto de la deuda externa pública y privada, etc.). En cambio, el metaanálisis recopila estudios efectuados por otros sobre el planteamiento del problema, los evalúa y puede volver a

analizar sus datos e integrarlos en una base mayor de información; o bien, adicionarlos y compararlos.

Krysik (2005) considera tres tipos de datos secundarios: datos obtenidos del nivel “micro”, datos agregados y datos cualitativos micro.

- a) *Datos obtenidos del nivel “micro”*: derivados de unidades de análisis individuales como personas, organizaciones u hogares (número de habitantes por vivienda, número de trabajadores de la empresa, preferencias electorales, satisfacción laboral de los trabajadores, etc.). La mayoría de las veces los datos se recolectan por medio de encuestas. Por ejemplo, los censos económicos (vistos en el capítulo de encuestas o *surveys* de este CD) generan esta clase de datos, al igual que los de población.

La encuesta nacional de adicciones (implementada anualmente en varios países latinoamericanos) es otro caso y recolecta —entre otros datos— los siguientes:⁹

1. Personas económicamente activas en el sistema familiar.
2. Distribución del tiempo libre del estudiante.
3. Convivencia del alumno en el hogar (padres y otros parientes).
4. Estructura del sistema matrimonial de los tutores del estudiante (padres casados, divorciados, separados, etcétera).
5. Ocupación de las personas económicamente activas en el hogar.
6. Persona(s) con quien(es) los estudiantes se comunican más dentro del hogar.
7. Actividades deportivas del estudiante (tipos y periodicidad de cada actividad deportiva, tiempos de dedicación, entre otros).
8. Conducta de fumar (inicio, frecuencia de la conducta de fumar, consumo diario de cigarrillos).
9. Ingesta de alcohol (inicio, periodicidad, consumo semanal).
10. Percepción proyectiva del consumo de estupefacientes o drogas ilegales (se responde a la pregunta: *¿A cuántos de tus compañeros de escuela has visto consumir droga?, ¿a cuántos de tus amigos, fuera de la escuela?*).

⁹ Por ejemplo, en México la publica periódicamente el Consejo Nacional Contra las Adicciones (CONADIC).

11. Conocimiento del tipo de droga que se consume (cuestionamiento: *Señala la droga que sabes consumen en este plantel... Señala la droga que sabes se consume en tu vecindario...*).
12. Preferencia de consumo en el entorno inmediato por tipo de estupefaciente (fuera de la escuela).
13. Percepción del encuestado respecto de si es un problema el consumo de estupefacientes.
14. Percepción respecto de si se pueden conseguir las drogas o estupefacientes dentro de la escuela.
15. Percepción respecto de si se consume droga o sustancias ilegales dentro de la escuela.
16. Percepción respecto de si el consumo de estupefacientes es un problema en su entorno externo.
17. Preferencia de consumo por tipo de estupefaciente en la escuela.
18. Actitud ante el consumo de estupefacientes por una amistad.
19. Manifestación de haber consumido algún estupefaciente.
20. Edad de inicio de consumo.
21. Figura que invitó a experimentar el consumo.
22. Tipo de estupefaciente o droga consumida por primera vez.
23. Consumo de diferentes clases de estupefacientes.
24. Preferencia individual de consumo por tipo de estupefaciente.
25. Periodicidad del consumo.
26. Motivo(s) del consumo.
27. Acciones recomendadas por los encuestados para reducir el consumo, evitarlo o prevenirlo.
28. Conciencia del consumo de estupefaciente.
29. Fuente más impactante de conocimiento del problema social que implica el consumo de estupefacientes.
30. Comentarios abiertos (cualitativo).

Pero el investigador, aunque no generó estos datos, tiene acceso directo a ellos (a la información individual y original; a los cuestionarios, registros, etc.) y, por lo tanto, los puede volver a analizar (incluso a veces puede identificar a los encuestados o sus perfiles).

b) *Datos agregados*: Consisten en datos individuales que ya se han procesado y conjuntado con otros en cifras estadísticas, tablas o cuadros. En éstos, las características de las unidades de análisis no pueden ser identificadas (por ejemplo: el investigador recibe el dato sobre el promedio de satisfacción laboral en una escala o la tasa de desempleo y obviamente no puede reanalizar los datos, sólo utilizarlos para los fines que convengan).

Estos datos se pueden obtener de:

1. Oficinas y agencias gubernamentales y oficiales (Ministerio de Educación, Ciencia y Tecnología de Argentina; Ministerio de Educación y Ciencia de España; Dirección de Impuestos y Aduanas Nacionales de Colombia; Instituto Nacional de Estadística de Bolivia; Ministerio de Industria, Turismo, Integración y Negociaciones de Perú; Banco Central de Venezuela; etcétera).
2. Organizaciones privadas (Cámara de Comercio de Santiago de Chile, Consejo de Fundaciones Privadas de Guatemala, Fundación Cultural Española para el Fomento de la Artesanía, Centro de Estudios Económicos del Sector Privado en México, Fundación Ford, entre otras).
Algunas bases de datos privadas contienen archivos periodísticos, como por ejemplo: U.S. News Archives on the Web (periódicos en Estados Unidos), Reuters y Europrensa (Universidad Complutense de Madrid), etc. Los periódicos más importantes del mundo generalmente poseen su propia base de datos.
3. Organizaciones no gubernamentales (Agencia de Cooperación en Ingeniería Ambiental o ACIA, Pro Amazonia "Selva Para la Vida", Alianza Cívica en México, Red 2002 de Barcelona para el apoyo de enfermos de SIDA, y otras).
4. Organizaciones internacionales (Comisión Económica para América Latina y el Caribe; Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, UNESCO; Organización Mundial de la Salud, OMS; Fondo de las Naciones Unidas para la Infancia, UNICEF; Fondo Monetario Internacional, FMI; etcétera).

Las primeras tres fuentes genéricas varían en cada país, por lo que resultaría impráctico intentar mencionarlas. Por ello, cada estudiante o profesor debe tener en mente cuáles son y qué tipo de información pueden proporcionar.¹⁰

De los datos resulta fundamental conocer:¹¹

- El nombre y la filiación institucional de quien recolectó los datos.
- El periodo de recopilación de los datos.
- La muestra (método de selección, tamaño, características...).
- El método de recolección de los datos, incluyendo el instrumento, procedimiento, confiabilidad y validez, etcétera.
- Localización (disponibilidad).

Y de ser posible:

- Procedimientos de codificación.
- Análisis practicados.
- Sugerencias para el análisis.

c) *Datos cualitativos micro*: producidos por una persona y recabados por otra. Al investigador le llega el dato (cartas recopiladas por otros, transcripciones de entrevistas hechas por psicólogos, grabaciones de las llamadas de pasajeros que iban en los aviones secuestrados por terroristas en los atentados a las torres gemelas del World Trade Center en Nueva York, etc.). Tales datos pueden volver a ser analizados por el investigador.

Las ventajas de los datos secundarios residen en que su costo es generalmente el más barato de todos los métodos de recolección y a veces son la única opción para el investigador (Smith, 2008; Sieppert, McMurtry y McClelland, 2005), además pueden replicarse estudios con base en ellos. Gran cantidad de datos secundarios se localiza mediante internet. Sus desventajas potenciales consisten en que a veces no son accesibles, no están estandarizados, pueden haberse omitido variables importantes para nuestro estudio y es factible que den una falsa impresión de ciertos subgrupos (los promedios de ingreso nacional per cápita, por ejemplo, suelen resultar engañosos).

¹⁰ Para el caso de España recomendamos a Berganza y García (2005) y a Corbetta (2003) y la siguiente página web: http://www.typicallyspanish.com/html/castellano/sociedad/gobierno/organizaciones_no_gubernamentales/. Para las organizaciones no gubernamentales latinoamericanas, por ejemplo, en el caso del medio ambiente, la siguiente dirección en internet incluye una gran cantidad de ellas: <http://www.ecoportal.net/content/view/full/158>

¹¹ Aplica también a los datos "micro".

El proceso para utilizar datos secundarios en una investigación se resume en la figura 7.14, adaptado de Sieppert et al. (2005) y Krysik (2005).

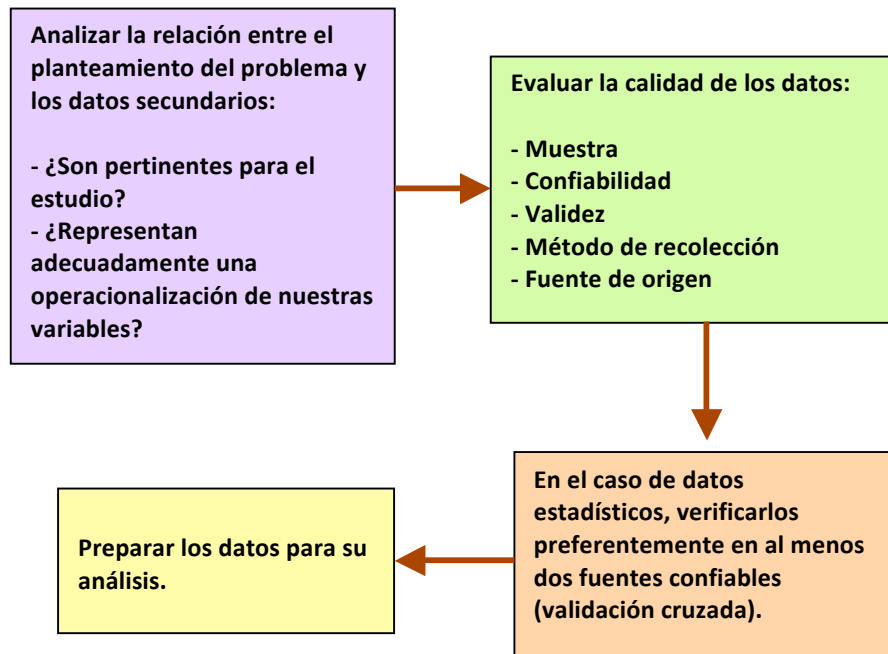


Figura 7.14 Proceso esencial para utilizar datos secundarios

Para el manejo de archivos se recomienda consultar a Webb, Campbell y Schwartz (2000).

El metaanálisis

Esta clase de procedimiento de recolección y análisis de los datos implica sintetizar resultados de investigación por medio de un número de estudios que consideran el mismo planteamiento y variables (Wiersma y Jurs, 2008).

Los pasos de un metaanálisis son muy similares a los de otros estudios, con la salvedad de que ahora la unidad de análisis ya *no* es el participante o caso, sino la investigación realizada por otro(s) investigador(es). El reto es resumir los datos a partir de los diferentes estudios de manera muy bien estructurada, que evite errores y facilite su ulterior análisis.

Cabe señalar que el metaanálisis será más preciso entre más homogéneos sean los estudios que se consideran, en cuanto a la pregunta de investigación, variables, método, instrumentos de medición y análisis estadístico. Por lo tanto, el primer paso consiste en ponderar si es posible combinarlos (nos referimos únicamente a estudios cuantitativos). Tal consideración nos puede llevar a un sinnúmero de opciones, desde descubrir que hay varios estudios que son totalmente compatibles y pueden mezclarse, hasta contrastar, aquellos que sólo tienen un ítem en sus cuestionarios que nos sea de utilidad.

Existen diferentes propuestas para conjuntar estadísticamente los resultados de distintos estudios, destacan tres fundamentalmente, que se conocen con el nombre de modelo de efectos fijos, modelo de efectos aleatorios y, por último, el modelo bayesiano (Molinero, 2003).

En el *modelo de efectos fijos*, los estudios se combinan y se asume que *no* existe heterogeneidad entre ellos, que por lo tanto todos constituyen estimaciones de un efecto real, cuya magnitud se desea conocer. Así pues, la inferencia realizada se condiciona a los estudios efectuados.

En el modelo de efectos aleatorios la inferencia se fundamenta en suponer que los estudios incluidos en el análisis constituyen una muestra aleatoria del universo de estudios posibles, y sus resultados son más conservadores al tener en cuenta una fuente extra de variación, ya que se incluyen dos fuentes de variación potenciales: la existente dentro de los estudios y la variación entre los estudios (Molinero, 2003, p. 3).

La utilización de modelos bayesianos requiere de modelos estadísticos complejos de varianza.

El metaanálisis constituye un intento de resolver contradicciones entre los resultados de diferentes investigaciones y evaluar patrones comunes en distintos estudios. Es un método sumamente importante para generar teoría.

Como todo método inferencial, tiene sus problemas y limitaciones. Uno de los principales problemas es que al ser un estudio que generalmente se ejecuta de forma retrospectiva, tiene una posibilidad mayor de introducir sesgos (error sistemático). Otros retos son los que se derivan del procedimiento de combinar datos de múltiples estudios, entre los cuales pudiera haber diferencias en cuanto a método y muestreo (Smith, 2008).

El procedimiento general de un metaanálisis es similar al de los datos secundarios, pero se agregan algunos elementos como el análisis mismo (figura 7.15).

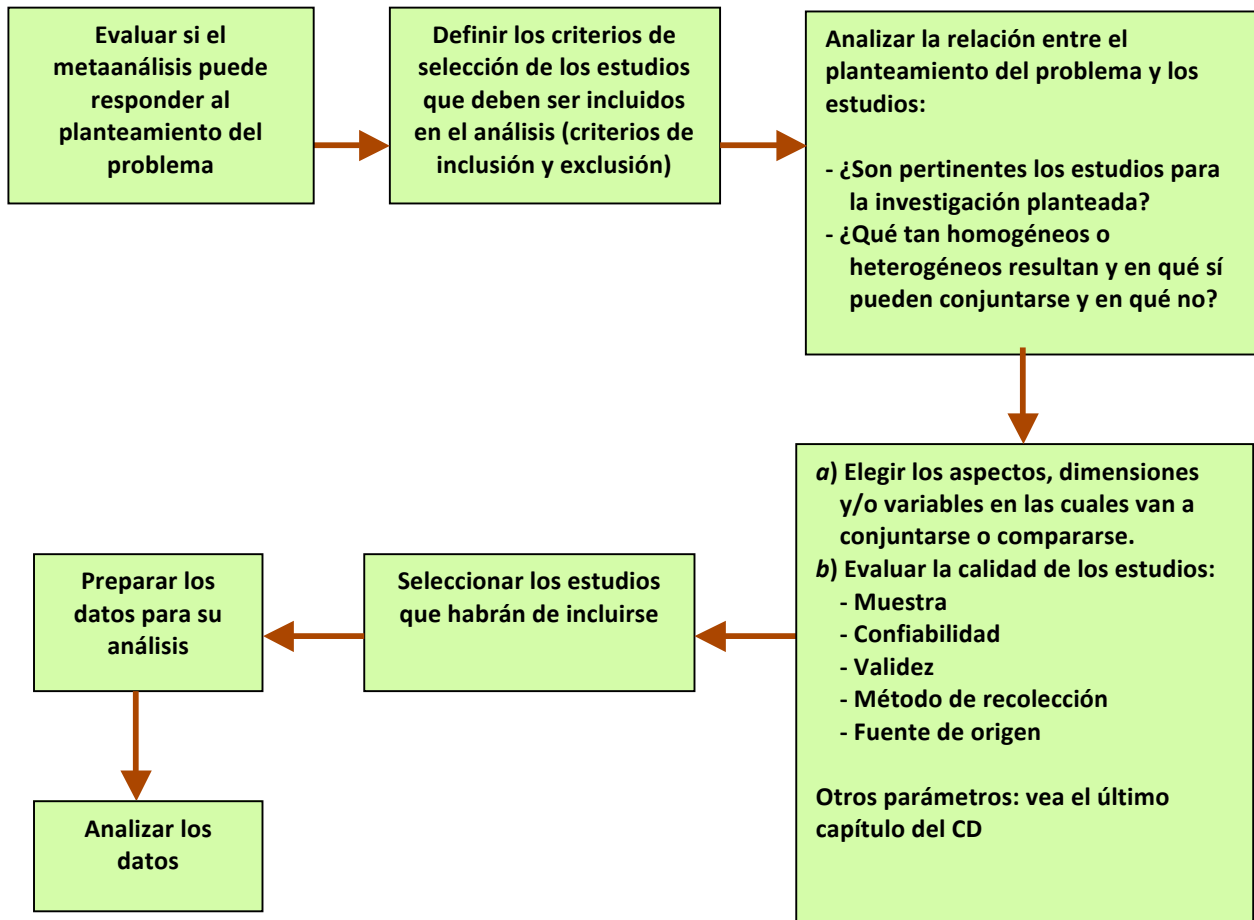


Figura 7.15 Proceso esencial del metaanálisis

Algunas recomendaciones para el metaanálisis las proporcionan Céspedes (1995) y Harpe (2009):

1. Seleccionar investigadores independientes que lean, clasifiquen, codifiquen, cuantifiquen y finalmente evalúen y elijan el grupo de estudios que serán incluidos en el metaanálisis (MA).

2. Elaborar una guía de los datos que deben ser recolectados de los estudios escogidos para ser procesados en el MA.
3. Combinar los resultados obtenidos y asegurar la calidad de los datos y su procesamiento estadístico adecuado.
4. Preferentemente utilizar datos provenientes de muestras aleatorias.

El análisis estadístico general para el modelo de efectos fijos (*no* existe heterogeneidad, hay homogeneidad o las investigaciones que se estiman muestran un efecto uniforme, debido al uso de un mismo tratamiento), se utiliza el método de Mantel-Haentzel-Peto.

Si los estudios son heterogéneos al mostrar el efecto ante un tratamiento (modelo de efectos aleatorios), se utiliza el método de Cochran-Dersimonian-Laird (Céspedes, 1995).

En general, se obtiene para las estadísticas calculadas, el tamaño del efecto (Nugent, 2009, Wiersma y Jurs, 2008, y Glass, 1976), el cual se revisó en el capítulo 10 del libro. Recordamos 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. Su cálculo fue ejemplificado con la prueba *t*.

Existen diversos programas computacionales para el metaanálisis, entre los cuales se encuentran los siguientes:

- Easy MA [M. Cucherat]. Lyon, France:
University of Lyon. <http://www.spc.univ-lyon1.fr/~mcu/easyima>
- RevMan®: The Cochrane Collaboration's Review Manager. The Cochrane Collaboration: <http://www.cc-ims.net/RevMan>
- BUGS y WinBUGS: <http://www.mrc-bsu.cam.ac.uk/bugs>
- SAS® (programa ya comentado, en su página existe un programa para metaanálisis). <http://www.sas.com>
- STATA®, <http://www.stata.com>
- Meta Win: <http://www.metawinsoft.com>

Y otros muchos que puede buscar vía disparador (o motor de búsqueda) en internet. Recuerde que las direcciones electrónicas pueden cambiar en cualquier momento.

EJEMPLOS DE ESTUDIOS BASADOS EN DATOS SECUNDARIOS

Shu-li et al. (2008) efectuaron una investigación para desarrollar un modelo que predijera la permanencia de personas ancianas en su hogar (en lugar de acudir a asilos, por ejemplo), utilizando datos secundarios de 9 879 individuos de la base de datos pública de Estados Unidos: “The National Home and Hospice Care Survey 2000 (NHHCS 2000)”. Descubrieron que el principal predictor de la permanencia es que la persona viva con familiares. Asimismo, Smith (2009) se fundamentó en análisis de datos secundarios para estudiar las inequidades en la educación británica, específicamente la exclusión de escuelas.

Hernández Sampieri et al. (2008) realizaron un estudio de mercado utilizando datos secundarios con la finalidad de indagar sobre la conveniencia de abrir clínicas de cirugía estética y plástica en México. Buscando en diversas fuentes de internet, medios impresos (revistas y diarios de circulación nacional) y bases de datos, lograron ubicar a las principales clínicas que operaban a nivel nacional y en las regiones y ciudades más pobladas del país (Ciudad de México, Monterrey, Guadalajara, Tijuana, Puebla), establecer los precios de sus servicios, determinar el mercado potencial y sus características (por ejemplo, número de cirugías que se practicaron durante el último año, interés por practicarse una cirugía plástica, preocupación por la estética personal...), evaluar las apelaciones de su publicidad y otras cuestiones. Haber realizado una encuesta hubiera sido más preciso, pero también más costoso, de tal modo que se hizo una investigación de acuerdo con el presupuesto del inversionista español. Incluso le proveyeron de un análisis del marco jurídico mexicano para abrir y operar tales clínicas.

Estudio para el diagnóstico municipal

En las anteriores ediciones de esta obra, se integraba un ejemplo para diagnosticar y caracterizar a un municipio, con el fin último de elaborar su plan de desarrollo. Regularmente estos diagnósticos son estudios mixtos, pero una parte muy importante de sus insumos lo constituyen los datos secundarios, que se generan con base en indicadores, los cuales se han desarrollado por diferentes organizaciones dedicadas a la investigación municipal (por ejemplo: el Instituto Vasco de Estadística; el Centro de Investigación y Desarrollo Económico de México (CIDE); el Instituto de Estadística de la Comunidad de Madrid; el Sistema Nacional de Indicadores Municipales (SINIM) de la Unidad de Información Municipal del Gobierno de

Chile; el Instituto Nacional de Estadística, Geografía e Informática en México (INEGI); el Instituto Brasileiro de Geografia e Estatística; el Sistema Nacional de Información Municipal, CEDEMUN, de la Secretaría de Gobernación de México; la Federación Nacional de Departamentos de Colombia; el Departamento de Estadística del Ayuntamiento de Barcelona; etcétera).

Algunos indicadores se aplican a cualquier municipio y son relevantes para todas las entidades poblacionales (número de habitantes, ingreso per cápita, ingresos totales del municipio, gasto total del municipio, inversión total, etc.), mientras que otros indicadores resultan importantes sólo para ciertos municipios (por ejemplo, los indicadores relativos a la producción agrícola son irrelevantes para un municipio completamente urbano). Por ello, en cada diagnóstico municipal deben elegirse los indicadores (datos secundarios) significativos para la entidad en cuestión.

También, cada provincia, departamento, región o estado genera indicadores a su nivel; lo mismo que las naciones.

En este CD: Material complementario→ Documentos→ Documento 3: “Indicadores”, se incluyen ejemplos de indicadores secundarios municipales que pudieran recolectarse para un diagnóstico de un municipio, con el fin de que los estudiantes que apenas comienzan en estos menesteres de la metodología de la investigación los puedan comprender sin mayores complicaciones (se cubren unas cuantas áreas, pero desde luego, no en su totalidad ni mucho menos). No son los únicos, ya que hay miles de ellos. Están agrupados por áreas generales (las cuales también pueden fluctuar según el autor u otros criterios). Determinados indicadores pueden formar parte de dos o más áreas. Sin polemizar (porque cuando se sopesan pueden generar discusiones severas), se presentan a continuación como muestras de datos secundarios y fueron en particular desarrollados para un municipio del Estado de México, México. Sería interesante como ejercicio que en cada institución se propusieran los que serían los más adecuados para su municipio.

Errores en las mediciones

Esta parte se escribe porque es una pregunta que se nos ha hecho constantemente sobre los tipos de errores en la medición.

En cualquier instrumento de medición se tiene un grado de error (recordemos lo señalado en el capítulo 9 del libro y al inicio de este capítulo).

Imaginemos que realizamos la evaluación de la autoestima en un grupo de estudiantes. Si la mido un día y la vuelvo a medir al día siguiente, y posteriormente a la semana; podemos esperar que las puntuaciones sean las mismas en las tres mediciones. La autoestima como otros atributos *no* varían a través del tiempo, a menos que “algo” los haga variar (un estímulo, la maduración de la persona, un suceso de vida, entre otros). Es decir, si los estudiantes obtuvieron puntuaciones diferentes en las pruebas, su desempeño se debió a la influencia de otro motivo distinto a su propia autoestima. Estas otras razones provocan error. El grado en que un instrumento se encuentra libre de error indica su confiabilidad. Entre más fiable resulta un instrumento, el investigador puede acercarse a una verdadera estimación del atributo considerado. Tal “exactitud” se obtiene al minimizar las fuentes de error en la medida de lo posible y con la estimación del grado de error que permanece. Pues bien, hay dos tipos de errores que pueden ocurrir: sistemáticos y no sistemáticos. Los primeros maximizan o minimizan el desempeño de un instrumento y afectan a la validez (Mertens, 2005). En el ejemplo de la autoestima, un método para desarrollarla o fortalecerla (por ejemplo, motivacional o terapéutico) puede ser visto como una influencia sistemática en la variable o atributo. El efecto del error sistemático en las respuestas es constante y puede predecirse. Pero los errores *no* sistemáticos (que conciernen a los investigadores) varían de situación en situación (aplicaciones distintas) y en consecuencia *no pueden predecirse* (o es sumamente complejo), pero *sí prevenirse*.

Estos errores pueden caer en tres categorías: *a)* personales (dependen de cada individuo, como el estrés, el cansancio o la motivación), *b)* condiciones de administración del instrumento (instrucciones distintas a los respondientes, temperatura ambiental, lenguaje, etc.) y *c)* modificaciones en los instrumentos o tareas (por ejemplo, cambios en ítems o las conductas a observar). Normalmente, la confiabilidad es calculada mediante una estadística que compara el desempeño de los mismos individuos en tiempos diferentes o partes distintas del instrumento y que oscila entre cero y uno, su fin es precisamente ayudar al investigador a identificar el grado de error que se presenta en la medición, en su consistencia interna; mediante la fiabilidad y un análisis del investigador se deben hacer transparentes las fuentes de error.