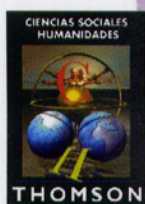
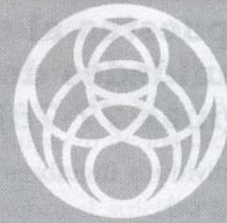


# Fundamentos de la investigación social



Earl Babbie



## *Conceptuación y medición*

---

### **Lo que aprenderá en este capítulo**

*Descubrirá que las palabras del lenguaje cotidiano comunican significados vagos e inespecíficos. En cambio, en la ciencia es esencial especificar exactamente lo que entendemos (y lo que no entendemos) con los términos que usamos.*

## Introducción

### La medición de todo lo que existe

¿Cómo se sabe?

Ideas y conceptos

Conceptuación

Indicadores y dimensiones

Carácter intercambiable de los indicadores

La confusión sobre definiciones y realidad

Creación de un orden conceptual

Ejemplo de conceptuación

## Criterios de las mediciones de calidad

Confiabilidad

Validez

¿Quién decide qué es válido?

Tensión entre confiabilidad y validez

## Puntos principales

## Preguntas y ejercicios de repaso

## Proyecto de continuidad

## Definiciones y propósitos de investigación

## Lecturas adicionales

---

## Introducción

Este capítulo es el primero de tres que se ocupan del proceso de pasar de las ideas vagas sobre lo que quiere estudiar a reconocerlo y medirlo en el mundo real. En este capítulo tratamos el tema general de la *conceptuación*, que dispone los cimientos para estudiar la *operacionalización* en el capítulo 6. En el capítulo 7, que examina las formas más complejas de medición, concluiremos los temas surgidos en los capítulos 5 y 6.

Quiero comenzar este capítulo con un ataque frontal a la preocupación que esconden algunas personas sobre la posibilidad de medir la sustancia de la vida: amor, odio, prejuicios, radicalismo, enajenación y asuntos por el estilo. La respuesta es sí, pero me tomará algunas páginas explicarme. Cuando usted vea que puede medir todo lo que existe, nos ocuparemos de las etapas necesarias para hacerla.

## La medición de todo lo que existe

Me parece totalmente posible que tenga algunas reservas sobre la capacidad de las ciencias para medir los aspectos verdaderamente importantes de la existencia social humana. Habrá leído informes de investigación que se ocupan de algo como el liberalismo o la religión o los prejuicios, y

tal vez se sintió insatisfecho con la forma en que los investigadores midieron lo que fuera que estudiaran. Acaso pensó que eran demasiado superficiales, que omitieron los aspectos que importan más. Quizá midieron la *religiosidad* como las veces que una persona iba a la iglesia o midieron el *liberalismo* de acuerdo con los votos de los ciudadanos en una sola elección. De seguro su insatisfacción habría aumentado si se descubriera mal clasificado por el sistema de medición. La gente suele tener esa experiencia.

O buscó en el diccionario la definición de una palabra como *compasivo*, la encontró deficiente y murmuró para sí mismo: "es más que eso". De hecho, cada vez que busca la definición de algo que ya entiende bien, es probable que imagine formas en que la gente confunda el término si sólo se basa en esa definición.

Ya hemos dicho que uno de los dos pilares de la ciencia es la *observación*. Como esta palabra puede sugerir una actividad más bien casual y pasiva, a menudo los científicos prefieren usar *medición*, con el significado de observaciones cuidadosas y deliberadas del mundo real, cuyo propósito es describir objetos y acontecimientos de acuerdo con los atributos que componen una variable. Si la variable en estudio fuera la *afiliación* a un partido *político*, consultaríamos el padrón de electores registrados para saber si las personas que estudiamos están registradas como

demócratas o republicanas. De esta manera habríamos medido su afiliación partidista. O en una entrevista que indague a *fondo* algunos temas políticos, simplemente les preguntáramos el partido con el que se identifican y tomaríamos su respuesta como nuestra medición. En general, existe más de una manera de medir una variable. Si existe, podemos medirla.

## ¿Cómo se sabe?

Para demostrarle que los científicos sociales pueden medir cualquier cosa que exista, quisiera que imagine que estamos debatiendo el asunto. Yo escribiré el guión, pero siéntase libre de modificar su parte del diálogo según le acomode.

YO: Los científicos sociales pueden medir todo lo que existe.

USTED: ¡Ja! Apuesto que no.

YO: Dígame algo que exista y le diré cómo medirlo.

USTED: Bueno, veamos su medición de los prejuicios.

YO: Buena elección. Pero no estoy dispuesto a perder nuestro tiempo tratando de medir algo que no existe, así que dígame si existen los prejuicios.

USTED: Claro que existen. Todos saben que los prejuicios existen, todos. Si usted fuera tan listo, también lo sabría. Hasta los tontos lo saben.

YO: Todos solían "saber" también que el mundo era plano. Quiero saber *cómo* sabe usted que los prejuicios existen en realidad.

USTED: Bueno, bueno. Como parece muy aferrado a las "observaciones", le diré que he visto los prejuicios.

YO: ¿Qué cosa ha visto usted que pruebe que los prejuicios existen?

USTED: Una vez, un empresario me dijo que nunca contrataría a una mujer para un puesto ejecutivo porque pensaba que todas eran caprichosas e irracional es. ¿Qué le parece?

YO: ¡Excelente! Eso me suena a prejuicios, así que supongo que podemos aceptar que existen. Ahora estoy preparado para medirlos. ¿Listo?

USTED: Listo.

YO: Usted y yo vamos a circular discretamente por las compañías locales para hablar con los empresarios acerca de las contrataciones. Cada vez que alguno nos diga que no contrataría a una mujer para un puesto ejecutivo porque piensa que todas son caprichosas e irracionales lo contaremos

como un caso de prejuicios. Cuando no se nos diga tal, lo anotaremos como un caso de desprejuicio. Al terminar, clasificaremos a los empresarios como prejuiciosos o desprejuiciados.

USTED: ¡Espere un momento! Ésa no es una medición muy buena de los prejuicios. Vamos a dejar fuera muchas clases. Nada más mediremos los prejuicios descarados en la contratación de mujeres. Justamente por eso dudo que pueda medirlos.

YO: Ya veo lo que quiere decir. Pero su comentario también significa que la situación que describió sólo prueba que existen los prejuicios descarados en la contratación de mujeres. Mejor reconsideremos si existen los prejuicios. ¿Existen?

USTED: Desde luego. Yo sólo le daba un ejemplo. Hay cientos de casos.

YO: Mencióneme uno que pruebe que los prejuicios existen.

USTED: De acuerdo, pruebe éste. Estaba yo en el campus la otra noche y dos tipos, uno blanco y otro negro, discutían sobre política. El tipo blanco se enojó tanto que comenzó a usar un espantoso lenguaje racista y gritó: "Habría que enviarlos a todos ustedes de regreso a África". ¿Es bastante prejuicioso para usted?

YO: Lo encuentro bien. Parece probar que existen los prejuicios, así que estoy listo para volver a medirlos. Nos dividiremos y pasearemos por el campus todas las noches. Iremos con las orejas bien abiertas hasta escuchar a un blanco que lance epítetos racistas y diga: "habría que enviarlos a todos..."

USTED: ¡Un momento! Ya sé adónde va y no es por ahí tampoco. Si alguien dice eso sería prejuicioso, pero clasificaríamos a mucha gente prejuiciosa como desprejuiciada sólo porque no se exalta ni usa epítetos racistas.

YO: Todo lo cual nos regresa a mi primera pregunta. ¿Existen realmente los prejuicios o sólo me está aguijoneando?

USTED: Sí, existen.

YO: Pues yo ya no estoy tan seguro. Usted me persuadió de que existen, empresarios que discriminan en el empleo a las mujeres; usted lo dijo y yo le creo. Me persuadió de que hay gente que insulta a los negros y les dice que regresen a África. Pero no estoy seguro de que existan *los prejuicios*. Desde luego que me gustaría encontrados para demostrarle que puedo medidos. Pero, para ser honesto, empiezo a dudar de que existan realmente; es decir, ¿alguna vez ha

visto prejuicios? ¿De qué color son?  
 ¿Cuánto pesan? ¿Dónde se localizan?  
 USTED: ¿De qué está hablando? ¿Hola? Tierra  
 llamando a Babbie. E. B., *phone home*.

El punto de este diálogo, como habrá conjeturado, es demostrar que *los prejuicios no existen*. No sabemos qué parecen, de qué tamaño o de qué color son. Nadie los ha tocado jamás ni ha cabalgado en ellos. Pero hablamos mucho de prejuicios. Veamos cómo sucedió.

Al recorrer el camino de la vida, usted y yo observamos muchas cosas y supimos que eran reales a partir de nuestras observaciones. Oímos muchas otras cosas que los demás nos dijeron que observaron y también parecían existir. Alguien dijo que vio un linchamiento y lo describió con lujo de detalles.

Conforme adquirimos más experiencia, advertimos algo más. Es muy probable que quienes participen en linchamientos también insulten a los negros. Más aún, muchos de ellos desean que las mujeres "se queden en su lugar". Al cabo, empezamos a comprender que había cierta clase de personas con esas tendencias. Cuando hablábamos de la gente que conocíamos, a veces resultaba conveniente identificar a alguien de acuerdo con dichas tendencias. Acostumbrábamos decir que fulano era "uno de esos que participan en linchamientos, insultan a los negros y no contratarían a una mujer para un puesto ejecutivo". Sin embargo, al poco rato se volvía engorroso decir todo eso y usted tuvo una idea brillante: "Vamos a usar la palabra *prejuicios* como abreviatura para referirnos a esas personas. Podemos usar el término aunque no hagan todas esas cosas, mientras se parezcan mucho".

Como soy complaciente y me interesa la eficiencia, accedí a emplear el sistema. De ahí salieron los *prejuicios*. En realidad nunca existieron, nunca los vimos. Sólo fue un atajo para hablar a espaldas de los demás. En última instancia, "prejuicios" no es más que un término que aceptamos usar en la comunicación: un nombre con el que representamos todo un conjunto de fenómenos aparentemente relacionados que hemos observado en el curso de la vida. Cada uno de nosotros nos formamos una imagen mental de lo que representan en general estos fenómenos reales que observamos, y de lo que tienen en común.

Cuando empleo la palabra *prejuicios*, sé que se activa una imagen mental en su mente, así como

en la mía. Es como si tuviéramos unos archiveros mentales llenos de miles de hojas con un rótulo en la esquina superior derecha. Una de las hojas en cada uno de sus archiveros tiene escrita la palabra *prejuicios*. Ahí se encuentra todo lo que le han dicho sobre los prejuicios y lo que usted ha observado que los ejemplifica. Mi hoja tiene lo que me han dicho más las situaciones que he visto que me parece que los ejemplifica.

## ***Ideas y conceptos***

El término técnico para estas imágenes mentales, estas hojas en nuestros archiveros mentales, es *idea*. Así, cada hoja es una idea. Ahora bien, no podemos comunicar directamente estas imágenes mentales. No hay modo de que pueda revelarle en forma directa lo que está escrito en las mías, así que nos valemos de los términos rotulados en la esquina superior derecha como medio de comunicar tanto nuestras ideas como lo que observamos que se les relaciona. Los términos vinculados a las ideas en nuestras mentes separadas nos posibilitan la comunicación y, en última instancia, concordar en lo que queremos decir específicamente con ellos. El proceso de ponerse de acuerdo es la *conceptuación*, y el resultado se llama *concepto*.

Supongamos que voy a conocer a alguien llamada Lola, a quien usted ya, conoce; por eso, le pregunto cómo es ella. Digamos que usted la ha visto ayudar a que los niños perdidos encuentren a sus padres y devolver pajaritos a sus nidos. Lola influyó en el hecho de que usted regalara pavos a las familias pobres el día de acción de gracias y en el de que visitara un hospital infantil en Navidad. La ha sorprendido llorando con una película sobre una madre que supera las adversidades para salvar y proteger a su hijo. Mientras repasa su archivero mental, encuentra que todos o casi todos estos fenómenos están anotados en una única hoja, etiquetada *compasiva*. Usted inspecciona otras entradas de la hoja y le parece que dan una descripción adecuada de Lola. Por tanto, me dice: "Lola es compasiva".

Ahora yo reviso mi archivero mental hasta que localizo una hoja titulada *compasiva*. Entonces, veo lo que está escrito y digo "ah, qué bien". Ahora creo que sé cómo es Lola, pero mis expectativas al respecto reflejan las entradas de *mi* archivero, no las del de usted. Después, cuando conozco a Lola, resulta que mis impresiones corresponden a las entradas de mí

## 100 Capítulo 5 Conceptuación y medición

hoja *compasiva* y le digo que tenía razón. Pero también comento que mis observaciones de Lola contradicen cosas que tengo en mi hoja y que no creo que sea muy compasiva. Comenzamos entonces a comparar notas.

Usted me dice que “una vez vi llorar a Lola con una película sobre una madre que vence la adversidad para salvar y proteger a su hijo”. Veo en mi hoja *compasiva* y no encuentro nada así. Buscando en otra parte del archivo, localizo esa clase de fenómeno en una hoja titulada *sentimental*. Replico: “Eso no es compasión; es sentimentalismo”.

Para fortalecer mi argumentación, le digo que vi que Lola se negó a dar dinero para una organización dedicada a salvar de la extinción a las ballenas. “Eso es una falta de compasión”, arguyo. Usted revisa su archivo y explica que tiene salvar a las ballenas en dos hojas: *activismo ambiental* y *encuentros entre especies*. Al cabo, nos ponemos a comparar las entradas de nuestras respectivas hojas tituladas *compasiva*. Descubrimos que las representamos con imágenes mentales muy diferentes.

En total, el lenguaje y la comunicación funcionan sólo porque las entradas de las hojas de nuestros respectivos archiveros se superponen considerablemente. Las similitudes de nuestras hojas representan los acuerdos de nuestra sociedad. De niños, nos dicen aproximadamente lo mismo cuando conocemos un término particular. Los diccionarios formalizan estos acuerdos que tiene nuestra sociedad sobre los términos. Así, cada uno de nosotros moldea sus imágenes mentales para que correspondan a dichos acuerdos, pero como tenemos experiencias y observaciones distintas, no hay dos personas que terminen exactamente con el mismo grupo de entradas en ninguna hoja de sus archiveros. Para regresar a la afirmación del principio del capítulo, podemos medir todo lo que es real. Por ejemplo, podemos medir si Lola de verdad regresa pajaritos a sus nidos, si visita el hospital en Navidad, si llora con las películas o si se niega a salvar a las ballenas. Todos estos fenómenos existen, así que podemos medirlos. ¿Pero en realidad es compasiva Lola? No podemos responder esta pregunta; en este sentido no podemos medir la compasión, porque no existe del mismo modo en que existen las cosas que acabamos de describir.

La *compasión* como *término* no existe. Podemos contar las letras y estar de acuerdo en que son nueve. Aceptamos que tiene tres sílabas y que comienza con C. En suma, podemos medir los aspectos que son reales.

Algunos aspectos de nuestras ideas también son reales. Tener o no una imagen mental del término *compasión* es real. Cuando un maestro de escuela primaria les pregunta a sus alumnos cuántos conocen el significado de *compasión*, es posible contar a los que levanten la mano. La presencia de determinadas entradas en las hojas rotuladas también es real y podemos medirla. Podemos medir cuántas personas relacionan el acto de dar dinero para salvar a las ballenas con su idea de la compasión. Lo único que no podemos medir es lo que sea realmente la compasión, porque no es real. La compasión existe sólo en la forma en que hemos acordado usar el término al comunicarnos cosas que son reales.

Si recuerda nuestro estudio del postmodernismo, imaginará que algunos objetarían el grado de “realidad” que confiero en los comentarios anteriores. En este capítulo no vamos a ser radicalmente posmodernos; creo que aceptará la importancia de una concepción intelectual dura de lo que es real y lo que no (cuando el intelectual se vuelve duro, se convierte en científico social).

En este contexto, Abraham Kaplan (1964) distingue tres clases de objetos que miden los científicos. La primera clase son los *observables directos*, aquello que observamos en forma simple y directa, como el color de una manzana o la marca en un cuestionario. Los *observables indirectos* requieren “observaciones relativamente más cuidadosas, complejas o indirectas” (1964:55). En un cuestionario, si observamos una marca en el recuadro *mujer*; hemos observado indirectamente el sexo de esta persona. Los libros de historia o las minutas de las juntas de los directivos de las empresas ofrecen observaciones indirectas de acciones sociales pasadas. Por último, los *constructos* son creaciones teóricas basadas en observaciones, que no pueden observarse en forma directa ni indirecta. El CI es un buen ejemplo. Es un constructo matemático de las observaciones de las respuestas a muchas preguntas en una prueba de inteligencia (en el capítulo 7 se encuentran otros ejemplos de constructos).

Kaplan (1964:49) define *concepto* como "una familia de ideas". Los conceptos -anota Kaplan- son constructos. Así, el concepto de la *compasión* es un constructo creado a partir de su idea de la compasión, de mi idea y de la de todos los que hayan empleado ese término. No se puede observar directa ni indirectamente porque no existe. Nosotros la construimos.

## Conceptuación

La comunicación diaria se establece mediante un sistema de acuerdos vagos y generales sobre el uso de los términos. Los demás no entienden exactamente lo que queremos comunicar, sino que captan el sentido general. Aunque usted y yo no estemos de acuerdo por completo en el uso del término *compasiva*, creo que puedo dar por hecho que Lola no le quitaría las alas a las moscas. El precio que pagamos por nuestra imprecisión es una gama amplia de malos entendidos y conflictos -entre personas y entre naciones-, pero de algún modo nos las arreglamos. En cambio, la ciencia aspira a algo más que salir del paso; no funciona en el contexto de tal imprecisión.

Catherine Marshall y Gretchen Rossman (1995:18) hablan de un "embudo conceptual" por el que los intereses de los investigadores se concentran cada vez más. Así, un interés general en el activismo social se restringiría a "los individuos dedicados al cambio social y a dar el poder a la gente", y aún más en descubrir "cuáles son las experiencias que originaron el desarrollo de activistas sociales completamente comprometidos". Este proceso de concentración se vincula en forma inevitable al lenguaje que utilizamos.

Como hemos visto, la *conceptuación* es el proceso por el que especificamos lo que queremos comunicar cuando usamos determinados términos. Por ejemplo, supongamos que queremos averiguar si las mujeres son más compasivas que los hombres. Sospecho que muchas personas dan por hecho que así es, pero sería interesante saber si así es en verdad. No podemos estudiar con provecho la cuestión, para no hablar de coincidir con el resultado, sin algún acuerdo de trabajo sobre el significado de *compasión*. Hablamos de acuerdos de *trabajo* en el sentido de que nos permiten trabajar en el tema. No tenemos que concordar -o pretender que concordamos- en que esa especificación sea en definitiva la mejor.

## Indicadores y dimensiones

El producto de este proceso de conceptualización es la especificación de uno o más *indicadores* de lo que tenemos en mente y que señalan la presencia o la ausencia del concepto que estudiamos. Veamos un ejemplo bastante sencillo.

Estaríamos, de acuerdo en que visitar hospitales infantiles en, Navidad y en Hanukkah es un indicador de compasión; devolver los pajaritos a sus nidos sería otro indicador, etc. Si la unidad de análisis de la investigación fuera el individuo, observaríamos entonces la presencia o la ausencia de cada indicador en cada persona de nuestro estudio. Aparte, sumaríamos los indicadores de compasión que observáramos en cada individuo. Por ejemplo, podríamos ponernos de acuerdo sobre 10 indicadores y encontrar seis en nuestro estudio de Lola, tres en Juan, nueve en María, etcétera.

Para regresar a nuestra pregunta original, digamos que calculamos que las mujeres de nuestro estudio tienen en promedio 6.5 indicadores de compasión mientras que los hombres tienen 3.2. Por tanto, concluiríamos sobre la base de nuestro análisis cuantitativo de las diferencias grupales que las mujeres son, en conjunto, más compasivas que los hombres. Sin embargo, por lo regular no es tan sencillo.

Imagine que a usted le interesa comprender un pequeño culto fundamentalista cristiano, en particular sus ásperas opiniones sobre varios grupos: homosexuales, ateos, feministas y otros. De hecho, sus miembros afirman que cualquiera que se niegue a unirse a su grupo arderá en el infierno. En el contexto de su antiguo interés en la compasión, no parecen tener mucha. No obstante, los escritos del grupo suelen hablar de su compasión por los demás.

Para explorar este interés de investigación, podría decidir alternar con los miembros del culto para conocerlos y aprender más sobre sus puntos de vista. Les diría que es un investigador social interesado en saber del grupo, o quizá sólo expresaría su interés sin decir por qué.

En el curso de sus conversaciones con los miembros y quizá en sus ceremonias religiosas, usted estaría en la posibilidad de llegar a entender lo que significa *compasión* para ellos. Por ejemplo, aprendería que están tan preocupados por los

pecadores que arden en el infierno que están dispuestos a ser agresivos, incluso violentos, para conseguir que la gente cambie sus costumbres pecaminosas. Por tanto, con este paradigma, consideran que es un acto de compasión golpear homosexuales, prostitutas y médicos que practican abortos.

Los científicos sociales suelen centrar su atención en los significados que las personas que son objeto de estudio dan a las palabras y los actos. Esto puede aclararles sus conductas, pero casi siempre complica los conceptos que les interesan.

Cuando analizamos con seriedad nuestros conceptos y especificamos lo que entendemos por ellos, descubrimos desacuerdos e inconsistencias. No sólo disentimos usted y yo, sino que además es probable que encontremos bastante confusas nuestras imágenes mentales. Si dedica un momento a considerar lo que *usted* entiende por compasión, verá que su imagen contiene varias *clases* de compasión. Las entradas en la hoja de su archivero se pueden componer en grupos y subgrupos, y hasta encontrará varios métodos para hacer combinaciones. Por ejemplo, podría agrupar las entradas en sentimientos y actos.

El término técnico para estos agrupamientos es **dimensión**: un aspecto o faceta especificable de un concepto. Así, hablaríamos de la "dimensión de sentimientos" de compasión y de la "dimensión de actos" de compasión. En otro agrupamiento, distinguiríamos compasión por los humanos" de "compasión por los animales". O podríamos pensar que la compasión consiste en ayudar a las personas a tener lo que *nosotros* queremos para ellas en lugar de lo que *ellas* quieren para sí mismas. Y todavía cabe distinguir la "compasión como perdón" de la "compasión como piedad".

Así, podemos dividir *compasión* en varias dimensiones. La conceptualización consiste tanto en especificar dimensiones como en identificar los indicadores de cada una.

A menudo, especificar las dimensiones de un concepto prepara el terreno para una comprensión más profunda de lo que estudiamos. Por ejemplo, podemos observar que las mujeres son más compasivas en términos de sentimientos y los hombres en términos de actos (o viceversa); pero, como quiera que sea, no podríamos decir si las mujeres son más compasivas que los hombres. De hecho, nuestra investigación habría mostrado que no hay una sola respuesta.

## Carácter intercambiable de los indicadores

Supongamos por un momento que usted y yo hemos compilado una lista de 100 indicadores de *compasión* con sus dimensiones. Digamos también que estamos muy en desacuerdo sobre los indicadores que prueban más claramente la presencia o la falta de compasión. Si estamos de acuerdo en algunos indicadores, podríamos centrar en ellos nuestra atención y es probable que coincidamos en la respuesta que dan. Pero supongamos que no concordamos en ningún indicador posible. Todavía podemos llegar a un acuerdo sobre si las mujeres o los hombres son más compasivos.

Si disentimos completamente en el valor de los indicadores, una solución sería estudiarlos todos. Supongamos que resulta que las mujeres son más compasivas que los hombres en los 100 indicadores (en todos los que usted favorece y en todos los míos). Entonces, podríamos convenir en que las mujeres son más compasivas aunque sigamos en desacuerdo sobre lo que signifique la compasión en general.

**El carácter intercambiable de los indicadores** significa que si varios indicadores representan en algún grado el mismo concepto, todos funcionarán de la misma manera que lo haría el concepto si fuera real y observable. Así, si en general las mujeres son más compasivas que los hombres, deberíamos observar esta diferencia utilizando cualquier medición razonable de la compasión.

Ya conoce la lógica fundamental de la conceptualización y la medición. Los análisis que siguen en este capítulo y el siguiente son ante todo perfeccionamientos y extensiones de lo que acabamos de presentar. Sin embargo, antes de pasar a las elaboraciones más técnicas del marco teórico principal, quiero tocar unos cuantos temas generales útiles.

Primero, sé que los párrafos anteriores quizá no se ajusten exactamente a su conocimiento del significado de términos como *prejuicios* y *compasión*. En la vida diaria nos conducimos como si tales términos tuvieran significados reales y definitivos. Por eso, en la siguiente subsección quiero comentar brevemente la manera en que llegamos a ese conocimiento.

Segundo, para que todo este análisis no produzca una imagen de anarquía en el significado de las palabras, vamos a describir algunas de las formas



en que los científicos sociales han organizado la confusión para normar y dar consistencia y generalidad al significado de los términos. Deberá terminar esta lectura con un nuevo sentido del orden, uno basado en la comprensión consciente más que en una aceptación casual del uso común.

## ***La confusión sobre definiciones y realidad***

Para empezar, un repaso breve. Los conceptos proceden de imágenes mentales (ideas) que resumen grupos de observaciones y experiencias que parecen estar relacionadas. Estas experiencias y observaciones son reales, por lo menos subjetivamente en tanto que los conceptos son meras creaciones mentales. Los términos asociados con los conceptos no son más que ardides creados con la finalidad de clasificar y comunicar. Por ejemplo, el término *prejuicios* es sólo un grupo de letras y no tiene significado intrínseco.

Ahora bien, solemos caer en la trampa de creer que los términos tienen significado real. El peligro aumenta cuando comenzamos a tomar en serio los términos y a tratar de usarlos con precisión. Más aún, el peligro es mayor en presencia de expertos que parecen saber más que uno sobre el significado real de los términos; en tal situación, es muy fácil ceder a la autoridad.

Una vez que hemos asumido que los términos tienen significados reales, iniciamos la tortuosa tarea de descubrir cuáles son los significados reales y qué constituye una medición genuina de ellos. La figura 5.1 ilustra la historia de este proceso. Fabricamos resúmenes conceptuales de observaciones reales porque son convenientes; pero esta conveniencia es tal que damos en pensar que son reales. Se llama **cosificación** al acto de considerar real aquello que no lo es. En la vida diaria, la cosificación de conceptos es bastante común.

El diseño y la realización de investigaciones sociales requiere que aclaremos la confusión entre conceptos y realidad. Para este fin, lógicos y científicos han encontrado útil distinguir tres clases de definiciones: *reales*, *nominales* y *operacionales*. La primera refleja la cosificación de los términos. Carl Hempel advierte:

De acuerdo con la lógica tradicional, una definición "real" no es una estipulación del significado de alguna expresión, sino una afirmación de la "naturaleza esencial" o los "atributos esenciales" de una entidad. Sin embargo, la noción de naturaleza esencial es tan vaga que esta caracterización es inservible para los propósitos de una investigación rigurosa.  
(1952:6)

La especificación de conceptos en la investigación científica depende de definiciones nominales y operacionales. Las definiciones *nominales* son las que se asignan a los términos. En medio de los desacuerdos y la confusión sobre lo que significa un término, podemos especificar una definición de trabajo para los propósitos de una investigación. Por ejemplo, si deseamos examinar en un estudio el nivel socioeconómico (NSE), podemos especificar simplemente que vamos a tratarlo como una combinación de ingresos y nivel de estudios. Con esta decisión, descartamos otros aspectos del NSE: ocupación, dinero en el banco, propiedades, linaje, estilo de vida, etcétera.

## ***Creación de un orden conceptual***

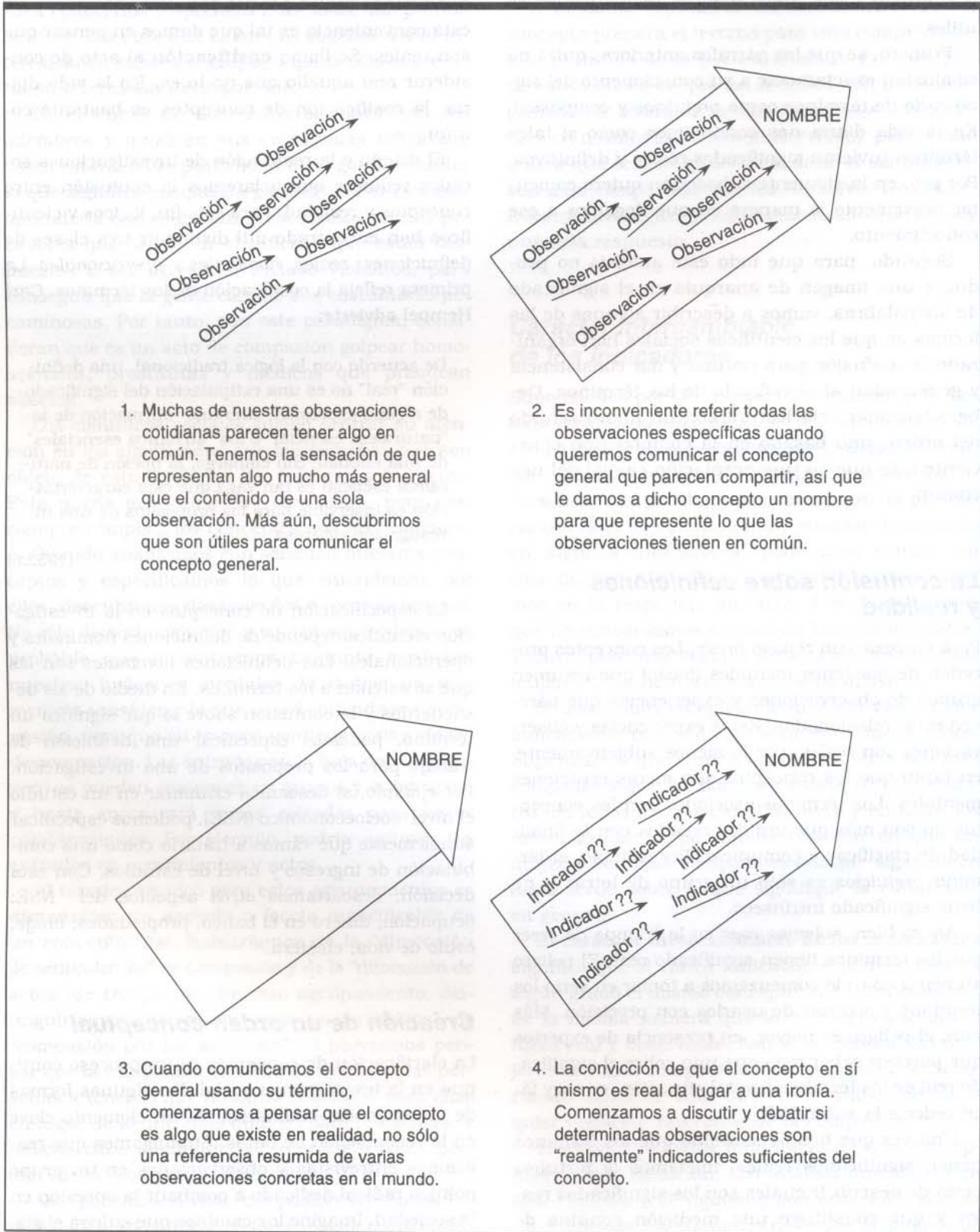
La clarificación de conceptos es un proceso continuo en la investigación social. En algunas formas de investigación cualitativa, es un elemento clave en la recopilación de datos. Supongamos que realizamos entrevistas y observaciones en un grupo político radical dedicado a combatir la opresión en la sociedad. Imagine los cambios que sufrirá el significado de *opresión* a medida que profundice en las experiencias y la visión del mundo de los miembros.

En el análisis de textos, los investigadores sociales hablan del "círculo hermenéutico", un proceso cíclico de comprensión cada vez más profundo.

La comprensión de un texto se da mediante un proceso en que el significado de las partes está determinado por el sentido general del texto tal como fue anticipado. En última instancia, la determinación más precisa de las partes puede cambiar el sentido anticipado del total, que de nuevo influirá en el sentido de las partes, y así sucesivamente.

(KVALE, 1996:47)

**Figura 5.1**  
El proceso de la confusión conceptual



El perfeccionamiento continuo de los conceptos ocurre en todos los métodos de investigación social. Usted verá que afina los significados incluso cuando escriba su informe final.

Aunque la conceptualización es un proceso continuo, es vital que uno la emprenda específicamente desde el comienzo de cualquier diseño de estudio, en particular en los que tienen una estructura rigurosa, como las encuestas y los experimentos. Por ejemplo, en una encuesta la operacionalización ocasiona una limitación al conjunto específico de preguntas del cuestionario que representan los conceptos que se estudian. Sin esta limitación, el estudio no avanzaría.

Sin embargo, también en los métodos de investigación menos estructurados se debe comenzar con un conjunto de significados anticipados que se puedan perfeccionar durante la recopilación e interpretación de los datos. Nadie cree de verdad que podamos observar la vida sin nociones preconcebidas; así, el observador científico debe estar consciente y ser explícito sobre estos puntos iniciales.

Exploremos la conceptualización inicial tal como se aplica a las investigaciones estructuradas, como las encuestas y los experimentos. El hecho de especificar las definiciones nominales concentra nuestra estrategia de observación, pero no nos permite observar. En la siguiente etapa debemos especificar exactamente qué es lo que vamos a observar, cómo lo haremos y qué interpretaciones asignaremos a las observaciones posibles. Todas estas especificaciones nuevas conforman la denominada **definición operacional** del concepto: una definición que explica con precisión la manera en que será medido. En estricto sentido, una definición operacional es una descripción de las "operaciones" que emprenderemos para medir el concepto.

Para continuar con el caso del NSE, podríamos formular a los entrevistados dos preguntas:

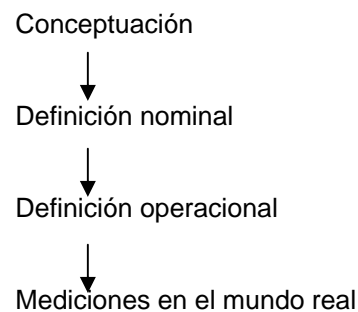
1. ¿Cuál fue el ingreso total de su familia durante los últimos 12 meses?
2. ¿Cuál es el mayor grado de estudios que cursó?

Aquí, tal vez quisiéramos especificar un sistema para clasificar las respuestas. En cuanto a los ingresos, estableceríamos categorías como "menos de 5000 dólares" o "de 5000 a 10 000 dólares". También agruparíamos en categorías el nivel escolar. Por último, especificaríamos cómo combinaríamos las respuestas de las dos

categorías para tener una medida del NSE (en el capítulo 7 presentaremos algunos métodos para lograrlo).

Al final, crearíamos una definición de trabajo viable del NSE. Aunque otros investigadores no estén de acuerdo con nuestra conceptualización y operacionalización, la definición tendría una virtud científica esencial: sería por completo específica e inequívoca. Aun en el caso de que alguien no concordara con nuestra definición, tendría una buena idea sobre cómo interpretar los resultados de nuestra investigación porque sería claro lo que entendemos por NSE (lo cual quedaría manifestado en nuestros análisis y conclusiones).

Veamos un diagrama de la progresión de las etapas de medición desde nuestra idea vaga de lo que significa el término hasta las mediciones específicas en una investigación científica completamente estructurada:



## Ejemplo de conceptualización

Quiero reunir los párrafos anteriores en una breve historia de un concepto social científico. Los investigadores que estudian las riñas callejeras suelen interesarse en el papel que juegan los sentimientos de impotencia. En este contexto, a veces emplean la palabra *anomia*. El gran sociólogo francés Emile Durkheim introdujo este término a las ciencias sociales en su clásico estudio de 1897: *El suicidio*.

A partir sólo de las publicaciones gubernamentales sobre los índices de suicidio en varias regiones y países, Durkheim produjo una obra de genio analítico. Para determinar los efectos de la religión en el suicidio, comparó los índices de los países predominantemente protestantes con los de las naciones predominantemente católicas, las regiones protestantes de los países católicos con las regiones católicas de los países protestantes,

etc. Para determinar los posibles efectos del clima, comparó los índices de naciones y regiones septentrionales y australes, y los examinó según los meses y las estaciones del año. Así, llegó a conclusiones sobre un acto supremamente individualista y personal sin tener dato alguno sobre las personas que lo cometieron.

En un nivel más general, Durkheim postuló que el suicidio también refleja el grado al que los acuerdos de la sociedad son claros y estables. Al observar que las épocas de agitaciones y cambios sociales enfrentan al individuo con graves incertidumbres sobre lo que se espera de él, Durkheim indicó que éstas le producen confusión, ansiedad e incluso lo llevan a destruirse. Para describir la condición social de falta de normas, eligió el término *anomia*. Vale la pena anotar que Durkheim no se sacó la palabra de la manga. Usada tanto en alemán como en francés, significa literalmente "sin ley", y el término en inglés, *anomia*, se emplea desde hace siglos con la significación peculiar de "desacato de la ley divina". Pero Durkheim hizo de *anomia* un concepto propio de las ciencias sociales.

En los años que siguieron a la publicación de *El suicidio*, los científicos sociales encontraron útil el término *anomia* como concepto y muchos ampliaron el significado que le asignó Durkheim. Robert Merton, en un artículo clásico titulado "Social Structure and Anomie" (1938), concluyó que la anomia es el resultado de una disparidad entre las metas y los medios prescritos por la sociedad. Por ejemplo, el éxito económico es una meta que muchos individuos comparten en nuestra sociedad, pero no todos cuentan con los recursos para alcanzarla por medios aceptables. Según Merton, resaltar la meta en sí provoca una falta de normas, porque los que rechazan las vías tradicionales para hacerse ricos proceden por medios ilegítimos para lograrlo. Por tanto, el estudio de Merton puede considerarse una conceptualización más profunda del concepto de anomia.

Aunque Durkheim tomaba el concepto de *anomia* como una característica de las sociedades, lo mismo que Merton después de él, otros científicos sociales lo han referido a los individuos (para aclarar la distinción, algunos estudiosos prefieren usar el término en su sentido social original y usar *anómico* para referirse a la característica del individuo). Así, en cualquier sociedad algunos individuos sufren de anomia y otros no. En un texto 20 años posterior al de Merton, Elwin Powell ofreció la siguiente conceptualización de anomia como característica de los individuos:

Quando los fines de la acción se vuelven contradictorios, inaccesibles o insignificantes, aparece una condición de anomia. Caracterizada por una pérdida general de orientación y acompañada de sentimientos de "vacío" y apatía. la anomia puede concebirse simplemente como una falta de significación.

(1957:132)

Powell indicó en seguida que hay dos clases de anomia y examinó cómo surgen en diferentes experiencias laborales y desembocan. a veces, en el suicidio. Sin embargo, Powell no midió la anomia *per se*, sino que estudió la relación entre suicidio y ocupación, e hizo inferencias sobre ambas clases de anomia. Así, el estudio no provee una definición operacional de anomia, sino sólo una nueva conceptualización.

Muchos investigadores han ofrecido definiciones operacionales, pero un nombre sobresale entre todos. Dos años antes de que apareciera el artículo de Powell, Leo Srole (1956) publicó un conjunto de reactivos para cuestionario que, según él, brindaba una buena medición de la anomia que sufren los individuos. Se trata de cinco enunciados con los que los sujetos deben o no estar de acuerdo:

1. A pesar de lo que digan algunos, la suerte del hombre común está empeorando.
2. No es justo traer niños al mundo dado cómo se ve el futuro.
3. En nuestros días, uno tiene que vivir para el hoy y dejar que el mañana se cuide solo.
4. En la actualidad, uno no sabe realmente con quién puede contar.
5. No tiene caso escribir a los funcionarios públicos porque no están interesados verdaderamente en el hombre común.

(1956:713)

En las décadas que siguieron a la publicación, la escala de Srole se convirtió en un elemento básico de la investigación para los científicos sociales. Descubrirá que esta operacionalización de anomia se emplea en muchos proyectos de investigación publicados en revistas académicas. De esto trata Srole en el recuadro, "Los orígenes del constructo 'anómico'", que preparó para este libro antes de su muerte.

En caso de que le interese en el concepto de *anómico*, he aquí unos datos de una encuesta realizada en 1994 en Estados Unidos por el Centro Nacional de Investigación de la Opinión de la Universidad de Chicago, como parte de la Encuesta

General Social permanente (véase la página en internet de la ESG en <http://www.icpsr.umich.edu/gss/subject/anomia.htm>). En cada una de las afirmaciones siguientes anoto el porcentaje de quienes están de acuerdo entre todos los que expresaron una opinión.

	Porcentaje de acuerdo
A pesar de lo que digan algunos, la suerte del hombre común no mejora, sino que empeora.	69%
No es justo traer niños al mundo dado cómo se ve el futuro.	45%
La mayoría de los funcionarios públicos No están interesados verdaderamente en los problemas del hombre común.	76%

Presenté esta historia muy abreviada de los conceptos sociales científicos de anomia y anómico por varias razones. Primera, porque ejemplifica el proceso mediante el cual los conceptos generales se convierten en mediciones operacionalizadas, aunque no quiero que piense que la distinción entre anomia y anómico está resuelta de una vez para siempre. Es seguro que los estudiosos los conceptuarán y operacionalizarán de nuevo en los próximos años, en busca de mediciones más útiles.

Por ahora, puse fin a la historia con la escala de Srole porque ilustra otro punto importante. El hecho de dejar abiertas la conceptualización y la operacionalización no produce necesariamente anarquía y caos, como acaso usted imaginaria. Con frecuencia surge un orden en el que figuran varios elementos. Primero, aunque uno *puede* definir anomia como quiera -digamos, la talla de zapatos-, lo más probable es que lo haga en formas que no difieran mucho de las imágenes mentales de los demás. Si usted adoptara una definición verdaderamente excéntrica, la gente lo ignoraría.

Segundo, cuando los investigadores descubren la *utilidad* de una conceptualización y operacionalización determinadas de un concepto, las adoptan, lo que origina definiciones uniformes de los conceptos. Además de la escala de Srole, citemos como ejemplos las pruebas de inteligencia y un caudal de mediciones demográficas y económicas que elaboran las oficinas de censos. El uso de mediciones establecidas tiene dos ventajas: han sido probadas y depuradas a fondo y es posible comparar los estudios que las utilizan. Si usted y yo realizamos sendos estudios sobre grupos distintos

y aplicamos la escala de Srole, podemos comparar nuestros grupos sobre la base de la anomia.

Por tanto, los científicos sociales pueden medir todo lo real, e incluso hacen un buen trabajo en la medición de lo que no lo es. Admitiendo que a fin de cuentas conceptos como el nivel socioeconómico, los prejuicios, la compasión y la anomia no son reales, vemos que los científicos sociales pueden crear un orden al manejarlos, un orden que no se basa en la veracidad, sino en la utilidad.

Dedicaremos el resto del capítulo a algunas consideraciones y alternativas para formular definiciones y mediciones útiles. Primero, veremos las relaciones entre las definiciones y los propósitos de las investigaciones. Después, concluiremos el capítulo con el examen de algunos criterios para determinar la *calidad* de las mediciones que concebimos.

## Definiciones y propósitos de investigación

Recuerde que en el capítulo 4 vimos que los dos propósitos generales de las investigaciones son la *descripción* y la *explicación*. La distinción entre ambas tiene implicaciones importantes para las definiciones y las mediciones. Si usted tiene la idea de que la descripción es más simple que la explicación, se sorprenderá de saber que las definiciones son más problemáticas en la investigación descriptiva que en la explicativa. Es necesario que sepa por qué sucede así antes de que nos ocupemos de otros aspectos de la medición, si bien en la parte 4 estudiaremos más a fondo el punto.

Es evidente la importancia de las definiciones para la investigación descriptiva. Si queremos describir e informar el índice de desempleo en una ciudad, nuestra definición de *estar desempleado* es crucial. Además, dependerá de nuestra definición de otro término: fuerza *laboral*. Si es del todo absurdo considerar desempleado a un niño de tres años, es porque no lo consideramos un miembro de la fuerza laboral. Así, en concordancia con una convención general de la oficina de censos, podemos excluir a los menores de 14 años de la fuerza laboral.

## Los orígenes del constructo "atómico"

por Leo Srole

Centro de Geriátría y Gerontología, Universidad de Columbia

**M**i larga fijación profesional en la anomia comenzó cuando leí *Le Suicide*, de Durkheim, mientras estudiaba en Harvard. Después, como alumno de posgrado en Chicago, estudié con dos antropólogos de la corriente de Durkheim: William Lloyd Warner y Alfred Radcliffe-Brown. Este último había mantenido una nutrida correspondencia con Durkheim, lo que me convierte en "descendiente" colateral del gran sociólogo francés.

Para mí, el primer impacto de la obra de Durkheim sobre el suicidio fue contradictorio pero permanente. Por un lado, tenía serias reservas sobre sus esfuerzos vigorosos, ingeniosos y a veces torpes por violentar los crudos registros burocráticos sobre tasas de suicidio para que casaran con su determinismo sociológico unidireccional. Por el otro, me conmovió su preocupación tenaz por la fuerza moral de los

lazos personales que nos vinculan a nuestro tiempo, lugar y pasado, y también su perspicacia al contemplar las consecuencias mortales que pueden seguir a la contracción y la descomposición de dichos lazos.

Mi interés en la anomia recibió una sacudida de testigo ocular al final de la Segunda Guerra Mundial, cuando trabajé para la Administración de Socorro y Rehabilitación de las Naciones Unidas (ASRNU) en la reconstrucción de una Europa despedazada por la guerra. En el campo de concentración de los nazis en Dachau vi de primera mano los abismos de deshumanización que las fuerzas macrosociales, como las que ocupaban a Durkheim, producían en individuos como Hitler, Eichmann y otros que obedecían sus dictados en todos los niveles de las fábricas nazis de la muerte.

Al volver de mi puesto en la ASRNU, sentía con toda urgencia que se acababa el tiempo para comprender la dinámica de la desintegración de los vínculos sociales. Necesitábamos trabajar en

Sin embargo, esta sola convención no nos da una definición satisfactoria, porque contaría como desempleados a los estudiantes de preparatoria, los jubilados, los incapacitados y las amas de casa. Podemos ampliar la convención y definir como miembros de la fuerza de labor a "todas las personas mayores de 14 años que trabajan, buscan trabajo o esperan ser reinstaladas en un puesto del que han sido despedidas o después de un permiso". Entonces, los desempleados serían todos los miembros de la fuerza laboral que no tienen empleo. Si un estudiante, una ama de casa o un jubilado no buscan trabajo, no debemos incluirlos en la fuerza laboral.

¿Pero qué significa "buscar trabajo"? ¿Debe uno estar registrado en una oficina oficial de colocaciones o debe ir de puerta en puerta pidiendo empleo? ¿Basta querer trabajar o estar disponible para una oferta de empleo? Convencionalmente, la definición operacional de "buscar trabajo" es responder *sí* al entrevistador que pregunta "¿ha buscado trabajo durante los últimos siete días?" (siete días es el periodo más especificado, pero

para algunos propósitos de investigación podría tener más sentido reducirlo o alargarlo).

He explicado con algún detalle estas consideraciones para que usted vea que la conclusión de un estudio descriptivo sobre, por ejemplo, el índice de desempleo depende directamente de cómo se resuelve cada cuestión. Incrementar el periodo que cuenta para decir que alguien busca trabajo tiene el efecto de añadir más desempleados a la fuerza laboral (tal como la definimos), y por tanto aumentar el índice de desempleo que se informa. Si seguimos otra convención y hablamos de fuerza laboral *civil* e índice de desempleo *civil*, excluiríamos a los militares, lo que también aumenta el índice de desempleo resultante, puesto que los militares estarían empleados *por definición*.

Así, la afirmación descriptiva de que el índice de desempleo en cierta ciudad es de tres por ciento o de nueve o de lo que sea depende en forma directa de las definiciones operacionales que se utilizaron. Este ejemplo es relativamente claro porque hay varias convenciones aceptadas en cuanto a la

forma expedita, aminorar la proliferación de las teorías más generales en favor del encuentro exploratorio directo con los individuos mediante la recién creada metodología moderna de la investigación con encuestas. También pensaba que una investigación así debía concentrarse en un espectro más amplio de patologías sociales que sólo el suicidio.

Mis primeras investigaciones fueron tentativas y diversas. Por ejemplo, en 1950 entrevisté a una muestra de 451 pasajeros de autobús en Springfield, Mass. Cuatro años después, el estudio de salud mental del centro de Manhattan me permitió llegar a una población mucho más grande. Estos proyectos de campo, entre otros, me proporcionaron la competencia para que expandiera y perfeccionara mis mediciones de la cualidad individual que reflejaba la cualidad macrosocial que Durkheim llamó *anomia*.

Al principio utilicé el término de Durkheim en mi propio trabajo, pero pronto decidí que era necesario limitar su uso a su significado

macrosocial y separarlo de tajo de sus manifestaciones individuales. Para este propósito, convenía adoptar la forma adjetiva del viejo término griego: *anómico* que se proponía solo.

El constructo del individuo *anómico* vio la luz en un artículo de 1956 en la *American Sociological Review*.<sup>\*</sup> Ahí expliqué las formas de operacionalizarlo y presenté los resultados de su primera aplicación en la investigación de campo. Para 1982, el índice de Referencias Científicas y el índice de Referencias de las Ciencias Sociales anotaba unas 400 publicaciones de ciencias políticas, psicología, trabajo social y sociología de todo el mundo que se habían valido de los instrumentos y descubrimientos de aquel artículo, lo que le valió que el Instituto Estadounidense de Información Científica lo designara "referencia clásica".

<sup>\*</sup>Leo Srole, "Social Integration and Certain Corollaries: An Exploratory Study", en *American Sociological Review* 21 (1956): 709-16.

fuerza laboral y el desempleo; pero piense en lo difícil que sería estar de acuerdo en las definiciones que necesitaríamos para decir que "45 por ciento de los estudiantes tienen posturas políticas conservadoras". Este porcentaje, como el índice de desempleo, depende directamente de la definición de lo que se mide. Otra definición propiciaría quizá la conclusión de que "cinco por ciento del estudiantado tiene posturas políticas conservadoras".

Irónicamente, las definiciones son menos problemáticas en el caso de la investigación explicativa. Supongamos que nos interesa explicar el conservadurismo político. ¿Por qué 'unas personas son conservadoras y otras no? Más en concreto, digamos que nos interesa saber si los ancianos son en general más conservadores que los jóvenes. ¿Qué ocurre si

usted y yo tenemos 25 definiciones operacionales de *conservador* y no podemos coincidir en cuál sea la mejor? Como ya vimos, éste no es un obstáculo necesariamente insalvable. Por ejemplo, supongamos que descubrimos que los ancianos son más conservadores que los jóvenes según *las 25 definiciones* (recuerde nuestro análisis de la compasión en hombres y mujeres). Supongamos, pues, que los ancianos son más conservadores que los jóvenes según *toda* definición razonable de conservadurismo que se nos ocurra. No importaría cuál fuera la definición; concluiríamos que los ancianos son en general más conservadores que los jóvenes aunque no nos hayamos puesto de acuerdo en lo que es el conservadurismo.

En la práctica, la investigación explicativa rara vez es tan inequívoca como en este ejemplo; no obstante, la pauta general es bastante común en la investigación real. *Existen* relaciones constantes en la vida social humana que dan por resultado hallazgos congruentes. Pero esta congruencia no aparece en una situación descriptiva. El hecho de cambiar las definiciones lleva casi inevitablemente a conclusiones descriptivas diferentes.

## Criterios de las mediciones de calidad

Hemos avanzado un poco en el capítulo. Comenzamos con la afirmación directa de que los científicos sociales pueden medir todo lo que existe. Entonces, descubrimos que casi todas las cosas que quisiéramos medir y estudiar en realidad no existen. Después aprendimos que de todos modos es posible medirlas. Quiero concluir el capítulo con un análisis de algunas de las varas con las que juzgamos nuestro relativo éxito o fracaso en la medición de las cosas, incluso de las que no existen.

Para comenzar, las mediciones se pueden realizar con varios grados de *precisión*, que representan la finura de las distinciones hechas entre los atributos que componen una variable. La descripción "una mujer de 43 años" es más precisa que "en sus cuarenta". Decir que una pandilla callejera se formó en el verano de 1996 es más preciso que decir "en la década de 1990".

Como regla general, las mediciones precisas son superiores a las imprecisas, como dicta el sentido común. No hay condiciones en las que las mediciones imprecisas sean intrínsecamente superiores a las precisas. Con todo, la precisión no siempre es necesaria o deseable. Si saber que una mujer está en sus cuarenta satisface los requisitos de su investigación, cualquier otro esfuerzo invertido en precisar su edad es un desperdicio. Así, la operacionalización de conceptos debe guiarse en parte por un conocimiento del grado de

precisión que se requiere. Si sus necesidades no están claras, sea más preciso y no menos.

Ahora bien, no confunda precisión con *exactitud*. Describir a alguien como "nacido en Stowe, Vermont" es más preciso que decir "nacido en Nueva Inglaterra"; pero supongamos que esta persona nació realmente en Boston. En este caso, la descripción menos precisa es un reflejo mejor y más exacto de la realidad.

Es evidente que la precisión y la exactitud son cualidades importantes en la medición científica y de hecho no requieren mayores explicaciones. Sin embargo, digamos aún que cuando los científicos sociales disponen y evalúan sus mediciones, prestan atención a dos consideraciones técnicas: **confiabilidad** y **validez**.

### **Confiabilidad**

En términos abstractos, la *confiabilidad* se refiere al hecho de que determinada técnica, aplicada repetidamente al mismo objeto, arroje cada vez el mismo resultado. Supongamos que usted quiere saber cuánto peso yo (quién sabe por qué quiere saberlo). Una técnica consistiría en pedirle a dos personas que estimen mi peso. Si una calcula 80 kilogramos y la otra 160, tendríamos que concluir que la técnica de hacer que la gente calcule mi peso no es muy confiable.

Supongamos como alternativa que usamos una báscula de baño como nuestra técnica de medición. Me subo dos veces a la báscula y anotamos ambos resultados. Es de creer que la báscula señalará el mismo peso las dos veces, lo que indicaría que es una técnica más confiable para medir el peso que pedirle a la gente que lo calcule.

Sin embargo, la confiabilidad no garantiza la exactitud más que la precisión. Supongamos que ajusté mi báscula para que me quitara dos kilogramos y medio, sólo para hacerme sentir mejor. Usted referiría el mismo peso las dos veces (confiabilidad), pero en ambas estaría



equivocado. En el capítulo 8 estudiamos este nuevo elemento, llamado **sesgo**. Por ahora, basta estar alerta al hecho de que la confiabilidad no garantiza la exactitud.

Supongamos que nos interesa estudiar la moral entre los obreros de dos clases de fábricas. En la primera, los obreros tienen puestos especializados de acuerdo con una división del trabajo muy acentuada. Cada obrero contribuye con una diminuta parte a la realización del proceso general en una línea de montaje larga. En la otra clase de fábricas, cada obrero desempeña muchas tareas y todo el proceso está a cargo de equipos de trabajo pequeños.

¿Cómo vamos a medir la *moral*? Si seguimos una estrategia, observaríamos a los obreros de cada fábrica para anotar si se hacen bromas, sonríen mucho, etc. Podríamos preguntarles si les gusta su trabajo e incluso si preferirían el sistema actual o el otro del estudio. Al comparar lo que observamos en las fábricas, llegaríamos a una conclusión sobre qué proceso de montaje eleva más la moral.

Ahora consideremos algunos problemas de confiabilidad inherentes a este método. Primero, nuestros sentimientos al realizar las observaciones de seguro matizarán lo que veamos. Tal vez interpretemos mal lo que veamos. Quizá los trabajadores bromeaban y creímos que estaban discutiendo o estaban en un día malo. Si durante varios días observáramos al mismo grupo de trabajadores, podríamos llegar a evaluaciones diferentes cada día. Por otro lado, si varios observadores evalúan la misma conducta, también podrían llegar a conclusiones distintas sobre la moral de los obreros.

Tomemos otra estrategia para evaluar la moral. Supongamos que verificamos los expedientes de la compañía para saber cuántas quejas se plantearon en el sindicato durante un período determinado. Es de creer que esto nos daría un indicador de la moral: entre más quejas, menos moral. Esta estrategia de medición parecería más confiable: si contamos las quejas una y otra vez, llegaríamos al mismo resultado.

Si al considerar la segunda estrategia usted de pronto piensa "¡un momento!", está preocupado por la *validez*, no por la confiabilidad. Terminemos nuestro análisis de la confiabilidad para poder ocuparnos de la validez.

En la investigación social, los problemas de confiabilidad aparecen de muchas formas. La

confiabilidad nos preocupa cada vez que un solo observador es la fuente de los datos, porque no tenemos ninguna salvaguarda contra el efecto de su subjetividad. No podemos estar seguros sobre qué parte del informe proviene de la situación observada y cuál del observador. Sin embargo, no es sólo un problema de los observadores únicos.

Quienes encuestan en sus investigaciones saben desde hace mucho que los entrevistadores, a causa de sus actitudes y conductas, reciben respuestas distintas de los entrevistados. Asimismo, si realizáramos un estudio de las posturas editoriales sobre algún asunto público, podríamos formar un equipo de codificadores que asumiera el trabajo de leer cientos de editoriales y clasificarlos de acuerdo con su opinión sobre el tema; pero codificarían de modo diverso el mismo editorial. También, si quisiéramos clasificar unos cientos de ocupaciones según un sistema fijo de codificación, digamos las categorías establecidas por el gobierno o el censo, usted y yo no colocaríamos las ocupaciones en las mismas categorías.

Todos estos ejemplos ilustran problemas de confiabilidad. Surgen problemas similares cada vez que le pedimos a las personas que nos den información sobre ellas mismas. Algunas veces formulamos preguntas que no saben cómo contestar: ¿cuántas veces ha ido a la iglesia? Otras, les preguntamos cosas que les parecen totalmente irrelevantes: ¿está satisfecho de las relaciones actuales entre China y Albania? Y otras más exploramos temas tan complicados que quienquiera que tenga una opinión clara al respecto podría interpretar de otra manera la pregunta si la formuláramos de nuevo.

¿Cómo elaborar mediciones confiables? Hay varias técnicas. Primera, al solicitarle información a la gente -si eso pide su diseño de investigación-, tenga cuidado de preguntar sólo lo que es probable que los entrevistados sepan contestar. Pregúnteles cosas importantes para ellos y hágalo en forma clara. En estos casos, el peligro es que la gente le *dará* respuestas, confiables o no. La gente le dirá lo que opina de las relaciones de China con Albania aunque no tenga ni la más vaga idea de cuáles sean.

Por fortuna, los investigadores sociales han concebido varias técnicas para enfrentar el problema básico de la confiabilidad.

**Método de test-retest** A veces es apropiado hacer la misma medición dos veces. Si usted no

contempla que la información cambie, entonces debe recibir la misma respuesta las dos ocasiones. Si las respuestas varían, el método de medición puede, en la medición de la variación, no ser confiable. Veamos un ejemplo.

En su investigación de la evaluación de riesgos para la salud (ERS), una parte de la medicina preventiva, Jeffrey Sacks, W. Mark Krushat y Jeffrey , Newman (1980) querían determinar el peligro de varios factores antecedentes y de estilo de vida, para que los médicos aconsejaran apropiadamente a sus pacientes. Si conocieran la vida de sus pacientes, los médicos podrían aconsejarlos sobre sus posibilidades de sobrevivencia y cómo mejorarlas. Desde luego, el objetivo dependía en buena medida de la exactitud de la información reunida sobre cada sujeto del estudio.

Para probar la confiabilidad de su información, Sacks y sus colaboradores pidieron a 207 sujetos que llenaran un cuestionario inicial sobre sus características y conducta. Tres meses después, los mismos sujetos respondieron un cuestionario de seguimiento sobre la misma información y se compararon los resultados de las dos encuestas. En general, apenas 15 por ciento de los sujetos dio la misma información en los dos estudios.

Sacks y sus colaboradores (1980:730) informaron lo siguiente:

Casi 10 por ciento de los sujetos señaló una estatura diferente en el segundo cuestionario. La edad de los padres cambió en uno de cada tres sujetos. Un padre envejeció 20 años en tres meses. Uno de cada cinco ex fumadores y ex alcohólicos tenía dificultades para recordar confiablemente su consumo anterior.

Algunos sujetos borrarón toda traza de soplos cardiacos, diabetes, enfisema, antecedentes penales e ideas de suicidio que habían revelado antes. La madre de un sujeto, muerta en el primer cuestionario, al parecer estaba viva y coleando en la época del segundo. A una de las encuestadas le faltaba un ovario en el primer estudio pero no en el segundo, mientras que otra a la que no le faltaba en el primer estudio ya no lo tenía en el segundo (ni en los 10 años anteriores). Un sujeto dijo que tenía 55 años en el primer estudio y 50 en el segundo, tres meses después. Uno se pregunta si los consejeros médicos llegarían a tener en sus pacientes siquiera parte del efecto que tenían sus

memorias. Así, el método de recopilación de datos no era especialmente confiable.

**Método de las mitades** Como regla general, siempre es bueno efectuar más de una medición de cualquier concepto social sutil o complicado, como los prejuicios, la enajenación o la clase social. Este proceder dispone los cimientos para otra verificación de la confiabilidad. Digamos que ha elaborado un cuestionario que contiene 10 preguntas que en su opinión miden los prejuicios contra las mujeres. Con la técnica de las mitades, usted dividiría al azar (véase el capítulo 8) las 10 preguntas en dos conjuntos de cinco. Como vimos en nuestra exposición del "carácter intercambiable de los indicadores" de Lazarsfeld, cada conjunto debe proporcionar una buena medición de los prejuicios contra las mujeres y ambos deben coincidir en la manera de clasificar a los sujetos del estudio. Si los dos conjuntos de preguntas miden a la gente en forma distinta, entonces aquí también hay un problema en la confiabilidad de la forma de medir la variable.

**Uso de mediciones establecidas** Otra forma de garantizar la confiabilidad al obtener información de la gente consiste en usar mediciones que han probado su confiabilidad en investigaciones anteriores. Por ejemplo, si quiere medir la anomia, tal vez convenga seguir los pasos de Srole.

Sin embargo, el uso frecuente de mediciones no garantiza su confiabilidad. Por ejemplo, las pruebas de aptitudes escolares y el inventario multifásico de personalidad de Minnesota (IMPM) han sido aceptados como normas establecidas en sus respectivos campos durante décadas. No obstante, en años recientes les ha hecho falta una revisión detallada para que reflejen los cambios de la sociedad.

**Confiabilidad de los asistentes** También es posible que la falta de confiabilidad proceda de los asistentes de la investigación, por ejemplo, entrevistadores y codificadores. Hay varias soluciones. Para precaverse de la falta de confiabilidad del entrevistador, es una costumbre común en las encuestas hacer que un supervisor llame por teléfono a una muestra de los entrevistados y verifique partes selectas de la información.

La repetición funciona también en otras situaciones. Si a usted le preocupa que las ocupaciones o los editoriales de los periódicos no estén bien clasificados, ¿por qué no hace que los codifiquen varios codificadores? Los que generen más

desacuerdos deben ser evaluados con más cuidado y resolverse en consecuencia.

Por último, la claridad, la especificidad, la capacitación y la práctica evitarán muchos problemas de falta de confiabilidad y fracasos. Si usted y yo dedicamos un tiempo a ponemos de acuerdo sobre cómo evaluar las posturas editoriales sobre cierta materia -si analizamos varias posturas y las leemos juntos con detenimiento-, es más probable que hagamos un buen trabajo al clasificarlos del mismo modo por nuestra parte.

La confiabilidad de las mediciones es un tema fundamental de la investigación social, y regresaremos a él más de una vez en los capítulos siguientes. Por ahora, recordemos que incluso la total confiabilidad no garantiza que nuestras mediciones midan lo que pensamos que miden. Estamos listos ya para sumergirnos en la cuestión de la validez.

## Validez

En el uso convencional, el término *validez* se refiere al grado en que una medición empírica refleja adecuadamente el *significado real* del concepto que consideramos. ¡Vaya! Ya nos habíamos adherido a la opinión de que los conceptos no tienen significado real. ¿Cómo, entonces, pretendemos llegar a saber si una medición refleja adecuadamente el significado del concepto? Desde luego, en última instancia no podemos. Al mismo tiempo, como vimos, toda la vida social, incluyendo la investigación social, opera según *consensos* sobre los términos que empleamos y los conceptos que representan. Hay varios criterios para determinar nuestro éxito en la medición adecuada de estos consensos.

Primero, existe algo llamado **validez patente**. Las mediciones empíricas pueden concordar o no con nuestros consensos y nuestras imágenes mentales respecto a cierto concepto. Usted y yo podríamos reñir por la suficiencia de medir la moral de los obreros contando las quejas presentadas en el sindicato, pero desde luego que coincidimos en que el número de quejas tiene que ver con la moral. Si propusiera que para medir la moral averiguáramos cuántos libros pidieron en la biblioteca los trabajadores en su tiempo libre, usted me haría una objeción más seria: esa medición no tendría mucha validez patente.

Segundo, ya señalé muchos de los consensos más formales y establecidos que definen algunos conceptos. Por ejemplo, el censo ha preparado definiciones operacionales de conceptos como familia, hogar y ocupación que tienen una validez práctica en la mayor parte de los estudios que los utilizan. Edward Carmines y Richard Zeller (1979) analizan otras tres formas de validez: *validez de criterio*, de *constructo* y de *contenido*.

**La validez de criterio**, a veces llamada *validez predictiva*, se basa en algún criterio externo. Por ejemplo, la validez de la junta escolar se muestra en su capacidad de predecir el éxito universitario de los estudiantes. En este sentido, la validez de una prueba escrita de manejo está determinada por la relación entre las calificaciones de la gente en la prueba y su historial posterior de conducción. En estos ejemplos, el éxito en la universidad y la capacidad de conducir son los *criterios*.

Para que ponga a prueba su comprensión de este concepto vea si se le ocurren conductas que sirvan para validar las siguientes actitudes:

Es muy religioso

Respalda por igual a hombres que a mujeres

Apoya a grupos paramilitares de extrema derecha

Está preocupado por el ambiente

A veces es difícil encontrar criterios conductuales para validar las mediciones en forma tan directa como en estos ejemplos. Sin embargo, en tales casos, podemos aproximar los criterios considerando cómo debería relacionarse teóricamente la variable en cuestión con otras variables. La **validez de constructo** se basa en las relaciones lógicas entre las variables.

Por ejemplo, supongamos que usted quiere estudiar la satisfacción matrimonial. sus causas y consecuencias, como parte del estudio concibe una medición de esta satisfacción y quiere evaluar su validez.

Además de idear su medición, también se formó ciertas expectativas teóricas sobre la relación de la variable *satisfacción matrimonial* con otras variables. Por ejemplo, podría concluir razonablemente que los cónyuges satisfechos tienden menos a la infidelidad que los insatisfechos. Si su medición relaciona la fidelidad matrimonial de la manera esperada, constituye una prueba de la validez de constructo de su medición. Sin embargo, si las parejas casadas satisfechas tendieran a ser

infiel tanto como las insatisfechas, se pondría en tela de juicio la validez de su medición.

Así, el examen de la validez de constructo confiere un *valor de prueba* de que su medición toca o no la cualidad que quiere medir, sin que sea una prueba definitiva. Si bien afirmo que las pruebas de la validez de constructo son menos sólidas que las de validez de criterio, advierta que hay lugar para los desacuerdos en cuanto a la clase de prueba que representa determinada variable en cierta situación. Menos importante que distinguir ambos tipos es entender la lógica de la validación que comparten: si hemos logrado medir alguna variable, entonces nuestras mediciones deben relacionarse de alguna forma lógica con otras mediciones.

Por último, la **validez de contenido** se refiere al grado en que la medición abarca la gama de significados que comprende el concepto. Por ejemplo, Carmines y Zeller (1979) señalan que un examen de las capacidades aritméticas no puede limitarse a la suma, sino que también debe cubrir la resta, la multiplicación, la división, etc. Asimismo, si decimos que medimos los prejuicios *en general*, ¿reflejan nuestras mediciones los prejuicios contra los grupos raciales y étnicos, las minorías religiosas, las mujeres, los ancianos y demás?

## ¿Quién decide qué es válido?

Al iniciar los comentarios anteriores sobre la validez recordé que dependemos de los consensos para determinar lo que es real, y acabamos de ver algunas de las formas en que los científicos sociales se ponen de acuerdo en que han hecho mediciones válidas. Hay todavía otra forma de considerar la validez.

A veces, los investigadores sociales se objetan unos a otros que asumen implícitamente que son superiores de alguna manera a quienes estudian. En efecto, a menudo buscamos motivaciones ocultas de las que no están al tanto los propios actores sociales. *Ustedes creen* que compraron ese Burpo-Blaster por su gran desempeño y buen aspecto, pero *nosotros sabemos* que en realidad lo que quieren es darse una posición social más elevada.

Este sentimiento implícito de superioridad cabe cómodamente en un planteamiento positivista a ultranza (el biólogo se siente superior a la rana de su mesa del laboratorio), pero choca con la postura más humanista y cualitativa que adoptan muchos científicos sociales. Así, por ejemplo, David

Silverman (1993:94-95) dice lo siguiente sobre la validez en el contexto de las entrevistas a fondo:

Si vamos a ver a los entrevistados como sujetos que construyen activamente las peculiaridades de su mundo cognoscitivo, entonces uno ha de lograr una intersubjetividad honda en ambos lados para llegar a una profunda comprensión mutua.

Al tratar de comprender la forma en que la gente común conceptúa e impone un sentido a su mundo, los etnometodólogos han pedido a todos los científicos sociales que respeten más estos procesos sociales naturales. Por lo menos, la conducta que acaso parezca irracional según el paradigma del científico puede verse lógica contemplada desde el paradigma del actor.

Como usted sabe, dediqué varios años al estudio de los trances mediumnámicos: el canalizador entra en trance y habla con una voz que dice que pertenece a otra persona, el espíritu de un muerto, un ser de otro planeta, etc. Desde el comienzo, vi con claridad que nunca le "encontraría el sentido" a lo que observara mientras exigiera que se acataran las reglas de la visión del mundo con que crecí. Así, por ejemplo, cuando sondeé el tema de "los buenos y los malos espíritus" quería descubrir los consensos entre los propios espíritus y entre quienes los canalizan.

En última instancia, los investigadores sociales deben considerar tanto a sus colegas como a sus sujetos fuentes de consenso sobre los significados y las mediciones más útiles de los conceptos que estudian. Algunas veces una fuente es más provechosa; otras, una distinta, pero ninguna debe quedar descartada.

## Tensión entre confiabilidad y validez

En una nota al pie de estas páginas, quiero señalar que suele haber una tensión entre los criterios de confiabilidad y validez, un estira y afloja. Si recuerda un momento el ejemplo de la medición de la moral en varias fábricas, creo que verá que la estrategia de sumergirse todos los días en la rutina de la línea de montaje, observar qué pasa y hablar con los obreros proporciona una medición más válida de la moral que contar las quejas. Parece obvio que uno se formaría una idea más clara de cuán alta o baja se encuentra la moral.

Sin embargo, como ya dije, la estrategia de contar sería más confiable. Esta situación manifiesta una tensión más general en las mediciones científicas.

Casi todos los conceptos realmente interesantes que queremos estudiar tienen muchos matices sutiles, y es difícil especificar con precisión lo que queremos que signifiquen. Los investigadores suelen decir que estos conceptos poseen una "riqueza de significados". Se han escrito muchísimos libros y artículos sobre el tema de la anomia, pero aún no lo han agotado.

Por tanto, con mucha frecuencia la especificación de mediciones y definiciones operacionales confiables parece robar a los conceptos su riqueza de significados. Después de todo, la moral es mucho más que la falta de quejas presentadas en el sindicato; la anomia es mucho más que los cinco enunciados de Leo Srole. Pero entre más toleremos variaciones y riqueza de significados en un concepto, más posibilidades hay de que no estemos de acuerdo sobre la manera en que se aplica en determinada situación, lo que reduce su confiabilidad.

En cierta medida, este dilema explica la persistencia de dos metodologías muy diferentes en la investigación social: técnicas cuantitativas, nomotéticas y estructuradas, como las encuestas y los cuestionarios, o los métodos cualitativos idiográficos, como las investigaciones de campo y los estudios históricos.

Para estar prevenido, debe guardarse de este dilema duradero e inevitable. Prepárese y enfréntelo. Si no hay un acuerdo claro sobre la forma de medir un concepto, mézalo de varias maneras. Si el concepto tiene varias dimensiones, mézalas todas. Por encima de todo, sepa que el concepto no tiene ningún otro significado aparte del que usted y yo le damos. La utilidad es la única justificación para asignarle cierto sentido a un concepto. Mida los conceptos en la forma que nos sirva para comprender el mundo que nos rodea.

---

## Puntos principales

- Las ideas son imágenes mentales personales que usamos como resúmenes para conjuntar experiencias y observaciones que parecen compartir algo. Empleamos términos o nombres para referirnos a esas ideas.
- Los conceptos son los significados aceptados que asignamos a los términos para facilitar la comunicación, la medición y la investigación.

- Nuestros conceptos no existen en el mundo real, así que no podemos medirlos directamente.
- Es posible medir lo que resumen nuestros conceptos.
- La conceptuación es el proceso de especificar las vagas imágenes mentales de nuestros conceptos y elegir las clases de observaciones y mediciones que serán apropiadas para nuestra investigación.
- El carácter intercambiable de los indicadores nos permite estudiar y sacar conclusiones sobre los conceptos aunque no estemos de acuerdo en sus definiciones.
- La precisión se refiere a la exactitud de la medición de una observación o la descripción de un atributo. Por ejemplo, describir a una persona como de "1.83 metros de estatura" es más preciso que decir "como de 1.80 metros".
- La confiabilidad se refiere a la probabilidad de que determinado procedimiento de medición arroje la misma descripción de cierto fenómeno si se repitiera. Por ejemplo, calcular la edad de una persona interrogando a sus amigos sería menos confiable que preguntar a la propia persona o consultar su acta de nacimiento.
- La validez es el grado al que cierta medición proporciona datos que se relacionen con los significados aceptados de un concepto. Existen varios medios de determinar la validez: validez patente, de criterio, de contenido y de constructo.
- La creación de mediciones específicas y confiables parece disminuir a veces la riqueza de significados que poseen nuestros conceptos generales. Se trata de un problema inevitable. La mejor solución es utilizar varias mediciones que exploren los diferentes aspectos del concepto.

---

## Preguntas y ejercicios de repaso

1. Tome un concepto de las ciencias sociales, como liberalismo o enajenación, y reespecifíquelo de modo que pueda estudiarse en un proyecto de investigación. No olvide especificar las dimensiones que quiere incluir y las que quiere excluir de su conceptuación.

2. Busque en un periódico o revista un caso de medición sin validez o confiabilidad. Justifique su elección.
3. Vaya a la página en internet de Estudios del Holocausto: Prejuicios (<http://www.socialstudies.com:80/12/126item.s.html>) y eche un vistazo a los materiales que se describen. Haga una lista de las dimensiones que se abordan de los prejuicios.
4. Examine las formas en que la Encuesta Social General ha explorado las opiniones sobre la pena de muerte (<http://www.icpsr.umich.edu/gss/subject/cuppun.htm>). ¿Qué aspectos del tema ha estudiado la ESG?

---

## Proyecto de continuidad

Hay muchas dimensiones en el concepto de igualdad y desigualdad sexual. Redacte una lista de por lo menos cinco dimensiones y proponga una forma de medidas. Está bien si utiliza diferentes técnicas de investigación para cada una.

---

## Lecturas adicionales

- Bohrnstedt, George W., "Measurement", en Peter H. Rossi, James D. Wright y Andy B. Anderson (comps.), *Handbook of Survey Research*, Nueva York, Academic Press, 1983, pp. 70-121. Este ensayo ofrece las bases lógicas y estadísticas de la confiabilidad y la validez en las mediciones.
- Carmines, Edward G., y Richard A. Zeller, *Reliability and Validity Assessment*, Beverly Hills, Sage, 1979. En este capítulo examinamos la lógica básica de la confiabilidad y la validez en las mediciones de las ciencias sociales. Carmines y Zeller exploran estos temas con más detalle y estudian algunas formas de calcular matemáticamente la confiabilidad.
- Gould, Julius, y William Kolb, *A Dictionary of the Social Sciences*, Nueva York, Free Press, 1964. Referencia fundamental de los acuerdos sociales científicos sobre varios conceptos. Aunque los términos que utilizan los científicos sociales no tienen en última instancia significados "verdaderos", este libro despliega los significados que tienen en mente cuando los emplean.

Grimes, Michael D., *Class in Twentieth-Century American Sociology: An Analysis of Theories and Measurement Strategies*, Nueva York, Praeger, 1991. Este libro ofrece un panorama excelente y de largo plazo sobre la conceptualización mientras el autor examina una variedad de posturas teóricas sobre las clases sociales y las técnicas de medición adecuadas.

Lazarsfeld, Paul F., y Morris Rosenberg (comps.), *The Language of Social Research*, Nueva York, Free Press of Glencoe, 1955, secc. 1. Excelente y variada colección clásica de descripciones de mediciones específicas en la investigación social del pasado. Estos 14 artículos son relaciones útiles y legibles de operaciones de medición reales efectuadas por científicos sociales, así como análisis más conceptuales de la medición en general.

Silverman, David, *Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction*, Newbury Park, Cal., Sage, 1993, cap. 7. Este capítulo aborda los temas de confiabilidad y validez específicamente en cuanto a la investigación cualitativa.

U.S. Department of Health and Human Services, *Survey Measurement of Drug Use*, Washington, D.C., Government Printing Office, 1992. Extenso repaso de las técnicas diseñadas y usadas para medir diversas formas de drogadicción.

Wallace, Walter, *The Logic of Science in Sociology*, Chicago, Aldine-Atherton, 1971, cap. 3. Presentación breve y lúcida de la formación de conceptos en el contexto de otras etapas de la investigación. Este análisis relaciona la conceptualización con la observación, por un lado, y con la generalización, por el otro.





## *Operacionalización*

---

### **Lo que aprenderá en este capítulo**

*Ahora pasaremos de la conceptualización a la siguiente etapa en la medición: ver de qué modo los científicos sociales encuentran los conceptos reflejados en el mundo real. En particular, consideraremos algunas destrezas que se requieren para formular preguntas.*



## En este capítulo...

---

### Introducción

#### Opciones de operacionalización

Margen de variación  
Variaciones entre extremos  
Nota sobre las dimensiones  
Niveles de medición  
Indicadores simples o múltiples

#### Algunos ejemplos de operacionalización

#### Lineamientos para formular preguntas

Preguntas y enunciados  
Preguntas abiertas y cerradas  
Formule reactivos claros  
Evite las preguntas dobles  
Los interrogados deben ser capaces de responder  
Los interrogados deben estar dispuestos a responder  
Las preguntas deben ser pertinentes

Los reactivos cortos son mejores  
Evite los reactivos negativos  
Evite reactivos y términos tendenciosos

#### Elaboración de cuestionarios

Formato general de los cuestionarios  
Formato para los entrevistados  
Preguntas condicionales  
Matriz de preguntas  
El orden de los reactivos del cuestionario  
Instrucciones  
Prueba previa del cuestionario  
Ejemplo compuesto

#### La operacionalización no cesa

#### Puntos principales

#### Preguntas y ejercicios de repaso Proyecto de continuidad

#### Lecturas adicionales

---

## Introducción

En el capítulo anterior analizamos y describimos diversos aspectos del proceso de conceptualización. En el curso de ese análisis hablamos con frecuencia de la operacionalización, puesto que ambas están íntimamente ligadas. Las distinguimos de esta *manera*: *conceptualización* es el perfeccionamiento y la especificación de conceptos abstractos, en tanto que la *operacionalización* es el desarrollo de los procedimientos concretos de investigación (operaciones) que darán por resultado observaciones de las representaciones de dichos conceptos en la realidad.

Mientras que las cuestiones de la medición que tocaremos en este capítulo son importantes para todas las formas de investigación social, muchas técnicas que veremos se relacionan de manera específica con la elaboración de cuestionarios, los cuales habitualmente se usan en las encuestas y los experimentos. En la parte 3 examinaremos las técnicas de operacionalización que se aplican a otros métodos.

Comenzaremos con un repaso de algunas opciones de operacionalización con que contamos a la hora de organizar el asunto de la observación y la medición: qué margen de variación considerar, qué niveles de medición emplear y si dependemos de un factor o de varios. Después, ejemplificaremos algunas formas de medir determinada variable, lo que espero que ensanchará su visión y su imaginación.

En seguida, debido a que la investigación social suele pedir información a la gente en encuestas, experimentos y estudios de campo, presentaremos algunos lineamientos generales y técnicas concretas para hacerlo en forma provechosa. Como veremos, hay muchos estilos de preguntas, de los cuales apenas unos cuantos arrojan información útil sobre el funcionamiento de la vida social humana. Un peligro de la observación es que nuestra lente de aumento se convierta en un espejo y sólo nos veamos a nosotros mismos.

Terminamos el capítulo con un examen de la operacionalización como un proceso que continúa durante todo el proyecto de investigación. Ya la

estudiamos en el contexto del diseño de investigación en lo que atañe a la recopilación de datos, y ahora veremos que los conceptos también se operacionalizan en el momento de analizar los datos. Estos párrafos finales redondearán su introducción a la medición en las ciencias sociales. En el capítulo 7 concluiremos el estudio de la medición.

## Opciones de operacionalización

Como dijimos, el investigador social dispone de una gran variedad de opciones cuando se trata de medir un concepto. Las más obvias son los métodos de recopilación de datos: encuestas, experimentos, investigación de campo, análisis de contenidos, investigación histórica, etc., que estudiaremos a fondo en la parte 3. Pero, como veremos en esta sección, los investigadores cuentan con muchas otras opciones. Aunque están vinculadas, las separaremos en aras de la exposición. Sin embargo, no olvide que la operacionalización no sigue una lista sistemática de verificación.

### *Margen de variación*

Al operacionalizar cualquier concepto, debe tener en claro el margen de variación que le interesa en su estudio. ¿Hasta qué grado está dispuesto a combinar atributos en categorías más bien generales?

Supongamos que desea medir los ingresos de las personas para un estudio después de reunir la información de expedientes o entrevistas. Los ingresos anuales más altos están en el orden de los millones de dólares, pero son muy pocos los que ganan tanto. A menos que estudie a los muy ricos, no sería muy útil formular y estar atento a categorías extremadamente altas. Dependiendo de los sujetos que estudie, lo más probable es que convenga establecer un piso mucho menor para la categoría de más altos ingresos: 100 000 dólares o más. Esta decisión lo llevará a colocar a personas que ganan billones de dólares al lado de otras que apenas obtienen 100 000, pero no les va a pasar nada, y la combinación tampoco dañará su investigación. En el otro extremo encarará la misma decisión. En estudios de población general, un corte de 5000 dólares o menos suele funcionar bien.

En el estudio de actitudes y tendencias, el tema de la variación tiene otra dimensión. Si no tiene cuidado, sin quererlo terminará midiendo la mitad del espectro de una actitud. Veamos un ejemplo de lo que pretendo decir.

Supongamos que le interesan las actitudes acerca del aumento en el uso de generadores nucleares de electricidad. De antemano, usted espera que algunas personas lo consideren el más grande avance desde la invención de la rueda, mientras que otras no tendrán en absoluto ningún interés en ello. Dada esta anticipación, parece que lo sensato es preguntar si se está en favor de ampliar el uso de la energía nuclear. Podría ofrecer categorías de respuesta que van de "Muy en favor" a "Nada en favor".

Sin embargo, esta operacionalización esconde la mitad del espectro de actitudes en cuanto a la energía nuclear. Muchas personas tienen una postura que no se limita simplemente a no estar a favor: se *oponen* completamente. En este caso hay una considerable variación a la izquierda del cero. Algunos se oponen un poco, algunos más y otros mucho. Entonces, para medir todas las variaciones, hay que operacionalizar las actitudes sobre la energía nuclear en un margen que vaya desde el punto de favorecerla mucho hasta el de no tener ninguna opinión en un sentido ni en otro, pasando por el de oponerse mucho.

Esta consideración es válida para muchas variables que estudian los científicos sociales. Prácticamente cualquier tema público tiene su apoyo y su oposición, ambas posturas con varios grados. Las tendencias políticas van de muy liberal a muy conservador, y, dependiendo del estudio, en ambos extremos hay lugar para los radicales. Las personas no son sólo más o menos religiosas; algunas son antirreligiosas.

No quiero decir que usted deba medir en todos los casos toda la gama de variaciones, sino que, según el propósito de su estudio, debe meditar si lo necesita. Si la diferencia entre *no religioso* y *antirreligioso* no es importante en su estudio, olvídelo. Alguien definió el pragmatismo como sigue: "cualquier diferencia que no marque una diferencia no es una diferencia". Sea pragmático.

Por último, su decisión sobre el margen de variación debe gobernarse por la distribución que espera de los atributos en sus sujetos de estudio. Esto es lo que quise decir arriba cuando afirmé que el margen depende de su objeto de estudio. En un estudio de las actitudes de los maestros

universitarios sobre el valor de la educación superior, nos detendríamos en *ningún valor* y no nos preocuparíamos por aquellos que la consideraran peligrosa para la salud de los estudiantes (pero si estudiáramos a los alumnos...).

## Variaciones entre extremos

En el capítulo 5 tratamos brevemente de la precisión como criterio de la calidad de las mediciones. También viene a cuento en la operacionalización de las variables. Todo se reduce a cuán finas serán sus distinciones entre los posibles atributos que componen una variable. ¿Importa si una persona tiene 17 o 18 años, o se podría realizar la investigación reuniendo un grupo de 10 a 19 años? No responda tan rápido. Si quiere estudiar los índices de empadronamiento y participación en las elecciones, en definitiva hay que saber si las personas estudiadas tienen edad para votar.

Entonces, si va a medir la edad, debe considerar el objetivo y los procedimientos de su estudio para decidir si las diferencias de edad pequeñas o grandes son importantes. Si mide las preferencias políticas, ¿importará si una persona es demócrata conservadora o demócrata liberal, o bien será suficiente conocer el partido? Al medir la afiliación religiosa, ¿basta saber que una persona es protestante o necesita conocer de qué iglesia? ¿Sólo necesita saber si una persona es casada o no, o bien hay diferencia en saber si nunca se ha casado o si es separada, viuda o divorciada?

Desde luego, no hay una respuesta general para estas preguntas. Las respuestas proceden del propósito de su estudio o del porqué realiza usted cierta medición. Sin embargo, puedo citar un lineamiento útil. Cuando no esté seguro de cuánto detallar una medición, opte por el exceso y no por el defecto. Durante el análisis de los datos puede combinar atributos precisados en categorías más generales, pero no puede dividir las variaciones que agrupó durante la observación y la medición.

## Nota sobre las dimensiones

Cuando ponemos manos a la obra de crear mediciones operacionales de las variables. a veces descubrimos o peor aún, nunca nos damos cuenta que no tenemos totalmente en claro qué dimensiones de una variable nos interesan. En el

capítulo 5 nos ocupamos del tema en alguna medida, pero ahora quiero que veamos más de cerca. Tomemos un ejemplo para ilustrar el punto.

Supongamos que usted y yo estudiamos las actitudes que tienen las personas acerca del gobierno y queremos incluir un examen de 10 que piensan de la corrupción. Veamos unas cuantas dimensiones que podríamos estudiar:

- ¿Piensan que hay corrupción en el gobierno?
- ¿Cuánta corrupción creen que hay?
- ¿Cuán seguras están de su juicio sobre el monto de la corrupción?
- ¿Qué piensan de la corrupción en el gobierno como problema de la sociedad?
- ¿Cuáles consideran que sean sus causas?
- ¿Les parece que es inevitable?
- ¿Qué creen que deba hacerse al respecto?
- ¿Qué están dispuestas a hacer para eliminar la corrupción en el gobierno?
- ¿Cuán convencidas están de que harían en efecto 10 que dicen?

La lista podría continuar indefinidamente; las actitudes hacia la corrupción en el gobierno tienen muchas dimensiones. Es importante saber con claridad cuáles son esenciales para la investigación; de otro modo, podríamos medir lo que opina la gente cuando en realidad queríamos saber cuánta corrupción creen que hay, o viceversa.

Cuando hayamos determinado cómo vamos a recopilar los datos (por ejemplo, encuestas. Investigación de campo) y decidido el margen de variación conveniente, el grado de precisión entre los extremos de la variación y las dimensiones concretas de las variables que nos interesan, nos aguarda aún otra elección, la elección lógico-matemática; es decir, hay que escoger el nivel de medición. Para examinar este asunto hemos de volver a los atributos y su relación con las variables (véase en el capítulo 1 el primer análisis del tema).

## Niveles de medición

Como recordará. los atributos son características o cualidades de algo. *Mujer* sería un ejemplo, así como *anciano* y *joven*. Por su parte, las variables son conjuntos lógicos de atributos. Así, *género* es

una variable compuesta de los atributos *mujer* y *hombre*.

Es posible considerar los procesos de conceptualización y operacionalización como la especificación de variables y los atributos que las componen. Así, en uno de los ejemplos del capítulo anterior, *ocupación* sería una variable con los atributos *empleado* y *desempleado*; la lista de atributos podría extenderse y comprender otras posibilidades analizadas.

Toda variable debe poseer dos cualidades importantes. Primero, los atributos que la componen deben ser *exhaustivos*. Para que una variable tenga alguna utilidad en la investigación, uno debe estar en posición de clasificar todas las observaciones según alguno de los atributos que componen la variable. Se meterá en problemas si conceptúa la variable *preferencias partidistas* de acuerdo con los términos *republicano* y *demócrata*, porque algunas de las personas que se dispone a estudiar se identificarán con el movimiento "Unidos resistimos", de Ross Perot, el Partido Verde o cualquier otra organización, y otras (muchas veces un gran porcentaje) le dirán que no tienen preferencias por ningún partido. Podría añadir *otros* o sin *preferencias* para hacer exhaustiva la lista de atributos. Como quiera que sea, debe ser capaz de clasificar *todas* las observaciones.

Al mismo tiempo, los atributos que componen la variable deben ser *mutuamente excluyentes*. Debe ser capaz de clasificar toda observación según *uno y sólo uno* de los atributos. Así, por ejemplo, necesita definir *empleado* y *desempleado* en forma tal que nadie pueda tener ambos atributos al mismo tiempo. Esto significa ser capaz de abarcar a la persona que tiene un empleo y busca trabajo (tal vez se tope con un mercenario de tiempo completo que aspira al *glamour* y la emoción de ser un investigador social). En este caso, podría definir sus atributos de manera que *empleado* tuviera precedencia sobre *desempleado*, con lo que cualquiera que tenga un trabajo está empleado, busque o no algo mejor.

Los atributos operacionalizados como exhaustivos y mutuamente excluyentes pueden relacionarse también de otra manera. Por ejemplo, quizá los atributos que componen las variables representan distintos *niveles* de *medición*. En esta sección examinaremos cuatro niveles de medición: nominal, ordinal, intervalar y proporcional o de razón.

**Mediciones nominales** Las variables cuyos atributos tienen sólo las características de exhaustividad y exclusión mutua son **mediciones nominales**. Entre los ejemplos se cuentan el *género*, *afiliación religiosa*, *preferencias partidistas*, *lugar de nacimiento*, *especialización universitaria* y *color de pelo*. Aunque los atributos que configuran estas variables *hombre* y *mujer* componen la variable *género* son distintos (y agotan las posibilidades del género), no tienen otras estructuras. Las mediciones nominales ofrecen nada más que nombres o títulos para las características.

Sería provechoso imaginar a un grupo de personas caracterizadas por una de tales variables reunidas físicamente según los atributos pertinentes. Suponga que se pide, en una reunión numerosa de personas, que se agrupen de acuerdo con el estado donde nacieron: en un grupo todos los que nacieron en Vermont; en otro, los que nacieron en California, etc. La variable sería *lugar de nacimiento*; los atributos, *nacido en California*, *nacido en Vermont*, etc. Todas las personas de determinado grupo tendrían por lo menos algo en común y diferirían de las personas en los otros grupos por la misma causa. Sería irrelevante saber dónde se formaron los grupos, qué tan cerca estaban o cómo se distribuyeron por la habitación. Todo lo que importaría sería que los miembros de cada grupo compartirían el mismo estado de nacimiento y que en cada uno sería diferente.

**Mediciones ordinales** Las variables con atributos que podemos disponer en un *orden lógico* son **mediciones ordinales**. Los distintos atributos representan más o menos cantidad de la variable. Las variables de este tipo son la clase social, el conservadurismo, la enajenación, los prejuicios, el nivel intelectual, etcétera.

En las ciencias físicas, la *dureza* es el ejemplo más citado de una medición ordinal. Decimos que un material (por ejemplo, el diamante) es más duro que otro (como el vidrio) si el primero puede rayar al segundo y no al contrario (es decir, el diamante raya al vidrio pero el vidrio no raya al diamante). Al tratar de rayar varios materiales con otros, al final seremos capaces de ordenarlos del más blando al más duro. Nunca determinaríamos la dureza de un material en términos absolutos, sino sólo relativos: qué materiales son más duros que los demás y cuáles más blandos.

Retomemos el ejemplo anterior en el que dividimos a los asistentes a una reunión e

## 122 Capítulo 6 Operacionalización

imaginemos que les pedimos que se coloquen en un grupo todos los que se titularon en la universidad, todos los que terminaron la preparatoria en otro y los que no la terminaron en otro más. Esta forma de agruparlos satisfaría los requerimientos de exhaustividad y exclusión mutua que ya señalamos. Pero además podríamos ordenar los grupos según el grado relativo de educación formal de cada uno (el atributo compartido). Así, los dispondríamos en fila de mayor a menor educación formal. Este ordenamiento proporcionaría una representación física de una medición ordinal. Si supiéramos en qué grupos están dos individuos, podríamos determinar que un individuo tiene más, menos o la misma educación formal que el otro. Del mismo modo, cualquier objeto puede ser clasificado como más duro, blando o igual que otro.

Es importante advertir que en este ejemplo sería irrelevante la distancia que separara a los agrupamientos por educación. El grupo de universitarios podría estar a 1.3 metros del grupo de preparatorianos y el grupo con menos que preparatoria a 130 metros. Estas distancias reales no tendrían ningún significado. Sin embargo, el grupo de preparatorianos tendría que estar entre el de universitarios y el otro, pues de otra manera el orden sería incorrecto.

**Mediciones intervalares** Ahora bien, en algunas variables es importante la distancia que separa a los atributos que la componen. Estas variables son **mediciones intervalares**: la distancia lógica entre los atributos puede expresarse en intervalos fijos significativos. Por ejemplo, en la escala de temperatura Fahrenheit la diferencia, o distancia, entre 80 y 90 grados es la misma que entre 40 y 50 grados Celsius; sin embargo, 80 grados Fahrenheit no es el doble de caliente que 40 grados Celsius, puesto que el punto cero de ambas escalas es arbitrario. Cero grados no significa falta de calor ni 30 representa 30 grados menos que cero calor (la escala Kelvin se basa en un cero absoluto, que en efecto significa ausencia de calor).

En general, las únicas mediciones intervalares habituales en las ciencias sociales son las construidas, como las pruebas estandarizadas de inteligencia que han sido más o menos aceptadas. El intervalo que separa las puntuaciones de CI de 100 y 110 puede tomarse como el mismo que separa las puntuaciones de 110 y 120 en virtud de la distribución de las calificaciones obtenidas por muchos miles de personas que han presentado

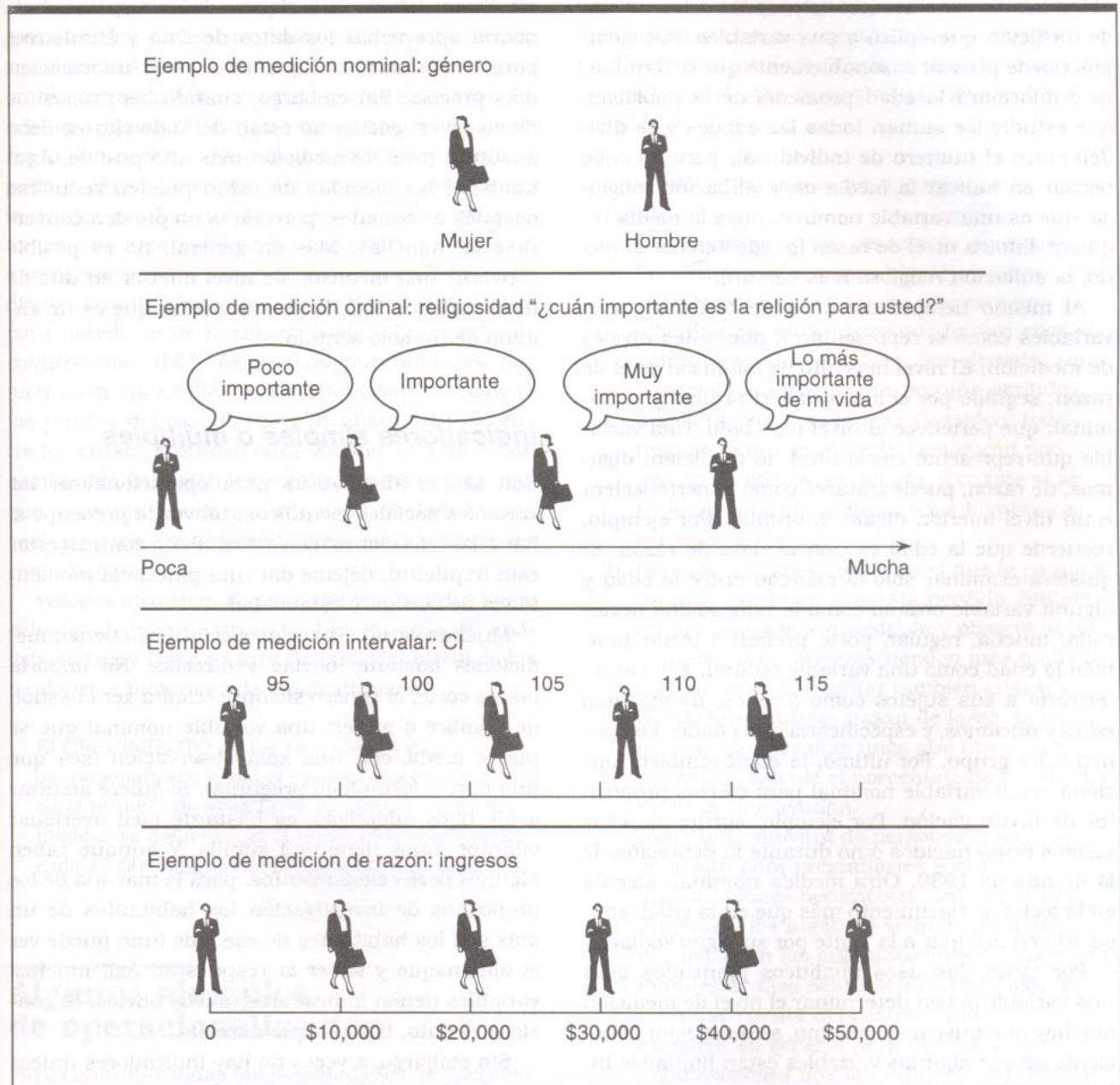
las pruebas durante años (si alguien obtuviera una calificación de cero en una prueba estandarizada de CI. Estrictamente hablando no la consideraríamos como falta de inteligencia, si bien nos parecería que no es adecuada para ser maestra universitaria y ni siquiera estudiante. ¿Pero qué tal uno de los decanos...?).

**Mediciones proporcionales o de razón** Casi todas las variables sociales científicas que cumplen con los requerimientos mínimos de las mediciones intercalares satisfacen también los de las **mediciones de razón**. En éstas, los atributos que componen las variables, aparte de las características estructurales ya señaladas, se basan en un punto cero verdadero. Ya señalamos la diferencia entre la escala de temperatura Kelvin y las escalas Fahrenheit y Celsius. Entre los ejemplos de la investigación social científica se encuentran *edad*, *tiempo de residencia en el mismo sitio*, *número de organizaciones de las que se forma parte*, *asistencias a la iglesia en cierto periodo*, *número de matrimonios* y *número de amigos árabes*.

De nuevo con el ejemplo de los juegos metodológicos en la reunión social, podríamos pedirles a los asistentes que se agruparan por edades. Todos los de un año estarían juntos (sentados o acostados), lo mismo los de dos años, de tres años, etc. El hecho de que los miembros de un grupo compartan la misma edad y de que cada grupo tenga una edad compartida diferente cumple los requisitos mínimos para una medición nominal. La disposición de los grupos en fila del más joven al mayor satisface los requerimientos adicionales de una medición ordinal y nos permite determinar si una persona es mayor, menor o de la misma edad que otra. Si separamos los grupos a distancias iguales, llenamos también los requisitos de la medición intervalar y podremos decir *cuán* mayor es una persona que otra. Por último, como uno de los atributos de la edad representa un cero verdadero (los bebés de madres que están a punto de alumbrar), la falange de infelices asistentes también cumple con los requisitos de una medición de razón, con lo que podemos decir que una persona tiene el doble de edad que otra.

Como repaso de este estudio, la figura 6-1 presenta una ilustración gráfica de los cuatro niveles de medición.

**Figura 6.1**  
Niveles de medición



**Implicaciones de los niveles de medición** Debido a que es poco probable que una empresa realmente los agrupe de personas que hemos descrito (inténtelo y no lo invitarán a muchas fiestas), quiero llamar su atención sobre algunas implicaciones prácticas de las diferencias que hemos señalado. En primer lugar, estas implicaciones aparecen en el análisis de datos (que

estudiamos en la parte 4), pero hay que anticiparlas desde la estructuración del proyecto de investigación.

Ciertas técnicas de análisis cuantitativo requieren variables que cumplan niveles mínimos de medición. Estas técnicas se deben planear en la medida en que las variables que examinará en su

proyecto se limiten a algún nivel de medición (digamos, ordinal). Más en concreto, debe prever que obtendrá las conclusiones adecuadas a los niveles de medición que aplicó a sus variables. Por ejemplo, puede planear razonablemente que determinará e informará la edad promedio de la población que estudie (se suman todas las edades y se dividen entre el número de individuos), pero no cabe pensar en indicar la media de la afiliación religiosa, que es una variable nominal, pues la media requiere datos a nivel de razón (puede señalar la moda, la afiliación religiosa más común).

Al mismo tiempo, usted puede tratar algunas variables como si representaran diferentes niveles de medición. El nivel más alto de medición es el de razón, seguido por el intervalar, el ordinal y el nominal, que pertenece al nivel más bajo. Una variable que represente cierto nivel de medición, digamos, de razón, puede tratarse como si perteneciera a un nivel inferior, digamos, ordinal. Por ejemplo, recuerde que la edad es una medida de razón. Si quisiera examinar sólo la relación entre la edad y alguna variable ordinal como la religiosidad declarada: mucha, regular, poca, preferiría tratar también la edad como una variable ordinal. Así, caracterizaría a sus sujetos como *jóvenes*, de *mediana edad* y *ancianos*, y especificaría qué edades componen cada grupo. Por último, la *edad* también funciona como variable nominal para ciertos propósitos de investigación. Por ejemplo, agruparía a los sujetos como nacidos o no durante la depresión de la década de 1930. Otra medida nominal, basada en la fecha de nacimiento más que en la edad, consistiría en agrupar a la gente por su signo zodiacal.

Por tanto, los usos analíticos planeados para una variable deben determinar el nivel de medición que hay que buscar. Asimismo, se debe estar consciente de que algunas variables están limitadas inherentemente a ciertos niveles. Si se va a emplear una variable de modo que requiere varios niveles de medición, hay que diseñar el estudio para llegar al nivel más alto. Por ejemplo, si se les pregunta a los sujetos del estudio su edad exacta, este dato puede organizarse después en agrupamientos ordinales o nominales.

Ahora bien, no tiene que medir por fuerza el mayor nivel de las variables. Si está seguro de que no necesitará un nivel mayor que el ordinal para las edades de las personas, puede pedirles que las indiquen por intervalos, como de 20 a 29, 30 a 39, etc. En un estudio de la riqueza de las corporaciones, podría aprovechar los datos de Dun y Bradstreet para ordenarlas en lugar de buscar

información más precisa. Sin embargo, cuando los propósitos de su investigación no están del todo claros, debe acudir al nivel de medición más alto posible. Aquí también, las medidas de razón pueden reducirse después a ordinales, pero éstas no pueden convertirse en aquéllas. Más en general, no es posible convertir una medición de nivel inferior en una de nivel superior. Vale la pena recordar que es un camino de un solo sentido.

## ***Indicadores simples o múltiples***

Con tantas alternativas para operacionalizar las variables sociales científicas, tal vez le preocupe si hará las elecciones correctas. Para contrarrestar esta inquietud, déjeme dar una pincelada momentánea de certeza y estabilidad.

Muchas variables sociales científicas tienen mediciones bastante obvias y directas. No importa dónde corte, el género siempre resulta ser cuestión de hombre o mujer: una variable nominal que se puede medir con una sola observación (sea que uno mire o formule la pregunta). Si quiere analizar a los hijos adoptivos, es bastante fácil averiguar cuántos niños tiene una familia. Y aunque caben algunos perfeccionamientos, para la mayoría de los propósitos de investigación los habitantes de un país son los habitantes de ese país (uno puede ver el almanaque y saber la respuesta). Así, muchas variables tienen indicadores únicos obvios. Si consigue el dato, tiene lo que necesita.

Sin embargo, a veces no hay indicadores únicos que den la medida de la variable que quiere. Como vimos en el capítulo 5, muchos conceptos están sujetos a interpretaciones variadas, cada una con varios indicadores posibles. En estos casos, conviene hacer varias observaciones de la misma variable. Entonces, se pueden reunir las unidades de información en una medición *compuesta* de la variable que interesa. Como todo el capítulo 7 está dedicado a los medios de hacerla, aquí veremos sólo un ejemplo sencillo.

Tomemos el concepto *desempeño en la universidad*. Todos hemos notado que algunos estudiantes se desenvuelven bien en la universidad en tanto que otros no tienen un desempeño tan bueno en sus cursos. Al estudiar el tema, podríamos preguntar qué características y experiencias se relacionan con los mejores desempeños; muchos investigadores lo han hecho. ¿Cómo debemos medir el desempeño general? La calificación de cada curso es un indicador posible del desempeño escolar, pero al tomar una sola

calificación corremos el riesgo de que no caracterice el desempeño general del estudiante. La solución es establecer firmemente lo que para usted, desde luego, es obvio: *el promedio de calificaciones* (PC). Asignamos puntuaciones numéricas a las calificaciones alfabéticas, sumamos los puntos del estudiante y dividimos el total entre los cursos seguidos para obtener una medición combinada (si varían los créditos de los cursos, hay que hacer ajustes en este respecto). En las ciencias sociales suele ser apropiado crear estas mediciones combinadas.

Veamos otro ejemplo de variables medidas como cálculos de más de una variable. Procede del análisis de Larry Isaac y Larry Griffin (1989:879) de las tendencias históricas de sindicalización.

El CRECIMIENTO de los sindicatos está operacionalizado como el cambio porcentual anual en el número de miembros. La frecuencia de las huelgas se define como el número de huelgas por cada 10 000 empleados no agrícolas participantes.

## Algunos ejemplos de operacionalización

Para conjuntar todas las posibilidades de operacionalización de que dispone el investigador social y mostrarle su alcance, quiero dedicar un espacio a ilustrar algunas formas de abordar ciertos problemas de investigación. Mi propósito es ampliar un poco más su imaginación y evidenciar el reto que la investigación social le plantea a su ingenio. Para simplificar el asunto, no describí todas las condiciones de la investigación que harían que una alternativa fuera mejor que las otras; pero usted debe estar al tanto de que no todas serán igualmente apropiadas en una situación dada. Veamos, pues, unas preguntas de investigación concretas y algunas maneras en que usted podría abordarlas. Comenzaremos con un ejemplo que expusimos a fondo en el capítulo 5. Tiene la ventaja adicional de que una de las variables es razonablemente sencilla.

1. *¿Son más compasivas las mujeres que los hombres?*
  - a. Seleccione un grupo de sujetos de estudio. Presénteles situaciones hipotéticas que conciernan a una persona que está en dificultades. Pregúnteles qué harían ellos si enfrentaran esa situación. Por ejemplo, ¿qué harían si se toparan con un niño perdido que llora por sus padres? Considere toda respuesta que manifieste compasión por ayudar al niño o consolarlo, y cuente si se inclinan más las mujeres o los hombres a ser compasivos.
  - b. Prepare un experimento en el que le pague a un niño por fingir que está perdido. Sitúelo en una banqueta transitada y observe si los hombres o las mujeres tienden más a ofrecer ayuda. No olvide contar también cuántos hombres y mujeres pasan de largo, puesto que pueden ser más unos que otras. Si así ocurriera, calcule el porcentaje de hombres y mujeres que ayudan.
  - c. Tome una muestra de personas y haga una encuesta para preguntarles a qué organizaciones pertenecen. Calcule si las mujeres o los hombres se inclinan más a participar en las organizaciones que parecen manifestar más sentimientos compasivos. Para dar cuenta del caso en que los hombres en general pertenezcan a más organizaciones que las mujeres o viceversa, haga lo siguiente: con cada persona de su estudio, calcule el *porcentaje* de su membresía a organizaciones que reflejen compasión. Vea si son las mujeres o los hombres quienes tienen el mayor porcentaje.
  - d. Busque en el periódico local algún artículo sobre un tema que tenga que ver con la compasión; por ejemplo, la matanza de crías de focas. En los días siguientes, lleve un registro de todas las cartas al editor sobre el asunto. Determine si son hombres o mujeres los que expresan más su compasión en la materia, y haga los ajustes necesarios si, en general, un género escribe más cartas que el otro.
2. *¿Qué lugar le parece a la gente el mejor para vivir; Nueva York o California?*
  - a. Consulte el *Statistical Abstract of the United States* o una publicación similar y verifique los índices de inmigración y emigración de cada estado. Trate de averiguar las cifras de los que se mudan directamente de Nueva York a California y viceversa.
  - b. Las compañías de sondeos de alcance nacional Gallup, Harris, Roper, etc., suelen preguntar a la gente cuál consideran el mejor estado para vivir. Busque algunos resultados recientes en la biblioteca o el periódico local.
  - c. Compare los índices de suicidio en los dos estados.
3. *¿Quiénes están mejor informados de los acontecimientos mundiales. los estudiantes de*



*sociología o los de contabilidad?*

- a. Prepare un cuestionario breve sobre acontecimientos mundiales y preséntelo a los estudiantes de una clase de sociología y a los de una clase equivalente de contabilidad. Si quiere comparar *especialidades* de sociología y contabilidad, asegúrese de preguntarle a los estudiantes apropiados.
- b. Pídale a un profesor de un curso de acontecimientos mundiales que le dé las calificaciones promedio de los estudiantes de sociología y de contabilidad que lleven su materia.
- c. Lleve a clases de sociología y contabilidad un pliego petitorio que solicite que "la sede de las Naciones Unidas se traslade a Nueva York". Cuento cuántos estudiantes de cada clase firman la petición y cuántos le informan que la sede de la ONU ya está en Nueva York.

Podríamos continuar de lo ridículo a lo más ridículo, pero el objetivo de estos ejemplos es ampliar su imagen de todas las variables que se pueden operacionalizar, no sugerir proyectos de investigación respetables. Si usted lo medita, absolutamente todo lo que ve a su alrededor ya es una medición operacionalizada de alguna variable. Casi todas son mediciones de más de una variable, así que todo lo que tiene que hacer es escoger las que prefiera y decidir qué representarán en su estudio. Por lo común, es mejor usar más de una medición de cada variable que se investiga.

---

## Lineamientos para formular preguntas

En los ejemplos anteriores y en la práctica real de la investigación social, las variables son operacionalizadas cuando los investigadores formulan preguntas a la gente como medio de recopilar datos para el análisis y la interpretación. Siempre ocurre así en las encuestas, y en los experimentos, las investigaciones de campo y otros modos de observación también se recaban datos de estos "informes personales". Algunas veces, un entrevistador hace las preguntas; otras, se dan por escrito a los interrogados (son los *cuestionarios de aplicación personal*).

Como los cuestionarios son un ejemplo común y concreto del proceso de operacionalización, son un tema que se presta para completar nuestro examen general. Como veremos, varios lineamientos generales le servirán para encuadrar y formular

preguntas que funjan como operacionalizaciones excelentes de las variables. También hay trampas que pueden dar por resultado información inútil y aun errónea; esta sección le ayudará a distinguirlas. Comencemos con algunas de las opciones en la elaboración de cuestionarios.

## Preguntas y enunciados

Aunque el término *cuestionario* sugiere un conjunto de preguntas, un examen de un cuestionario común revelará tantos enunciados como preguntas. Esto no carece de razón. A menudo, al investigador le interesa determinar hasta qué grado los entrevistados sostienen cierta actitud o punto de vista. Si uno puede resumir tal actitud en un enunciado breve, se puede preguntar a los interrogados si están de acuerdo o no lo están. Rensis Likert formalizó en buena medida este procedimiento con la creación de la escala que lleva su nombre, un formato en el que se pide a los entrevistados que estén muy de acuerdo, de acuerdo, en desacuerdo o muy en desacuerdo, o bien que aprueben mucho, aprueben, etcétera.

Es posible sacar provecho tanto de preguntas como de enunciados. El uso de ambos en un cuestionario da más flexibilidad para redactar los reactivos y también puede hacerlo más interesante.

## Preguntas abiertas y cerradas

Los investigadores tienen dos opciones al plantear preguntas. Así, pueden formular *preguntas abiertas*, caso en el que el se pide al **interrogado** que dé su propia respuesta. Por ejemplo, se puede preguntar al interrogado "¿cuál cree que es el asunto más importante que enfrenta actualmente el país?" y dejar un espacio para la respuesta (o que responda de palabra a un entrevistador).

En el otro caso, las *preguntas cerradas*, se pide al entrevistado que elija una respuesta de la lista que le presenta el investigador. Las preguntas cerradas son muy populares porque ofrecen respuestas más uniformes y se procesan con más facilidad. Las respuestas a preguntas abiertas tienen que codificarse para procesarlas en el análisis por computadora, como veremos en el capítulo 14. Con frecuencia, esta codificación requiere que el investigador interprete el sentido de las respuestas, lo que abre la posibilidad de errores de comprensión y sesgo. También se corre el peligro

de que los interrogados den respuestas esencialmente irrelevantes para las intenciones del investigador. En cambio, las respuestas a preguntas cerradas se transfieren por lo regular directamente al formato de computadora.

El principal inconveniente de las respuestas a preguntas cerradas radica en la estructuración de las respuestas por parte del investigador. Cuando las respuestas pertinentes a cierta pregunta son relativamente claras, no habrá problemas. Sin embargo, en otros casos la estructuración del investigador puede ignorar algunas respuestas importantes. Por ejemplo, al preguntar sobre "el asunto más importante que enfrenta actualmente el país", su lista de verificación de los temas puede omitir ciertas cuestiones que en opinión de los entrevistados sean importantes.

La redacción de preguntas cerradas debe guiarse por los dos requisitos estructurales que ya estudiamos. Las categorías de las respuestas ofrecidas deben ser *exhaustivas*: deben comprender todas las respuestas posibles que se anticipen. A menudo, los investigadores lo garantizan añadiendo una categoría llamada, digamos, Otros (especifique:\_\_\_\_\_).

Segundo, las categorías de las respuestas deben ser *mutuamente excluyentes*. El interrogado no debe sentirse inclinado a elegir más de una (en algunos casos, querrá seleccionar varias respuestas, pero esto traería dificultades posteriores para el procesamiento y el análisis de los datos). Para asegurarse de que sus categorías sean mutuamente excluyentes, medite con cuidado en todas las combinaciones de las categorías y pregúntese si cabe la posibilidad de que alguien escoja más de una respuesta. Además, es útil añadir instrucciones que indiquen a los entrevistados que elijan la mejor respuesta, pero esta técnica no es un sustituto satisfactorio de las respuestas elaboradas cuidadosamente.

### **Formule reactivos claros**

No debería ser necesario decir que los reactivos del cuestionario deben ser claros e inequívocos, pero la enorme proliferación de preguntas oscuras y ambiguas en las encuestas hace que valga la pena subrayar el punto aquí. Muchas veces, uno está tan absorto en el tema de la investigación que las opiniones y los puntos de vista le parecen claros, pero no a los entrevistados (de los que muchos prestarán poca o ninguna atención al tema). O si

uno tiene conocimientos apenas superficiales del tema, puede no ser capaz de especificar en forma suficiente el sentido de la pregunta. La pregunta "¿qué opina del plan de paz propuesto?" puede suscitar en el interrogado una réplica: "¿cuál plan de paz?". Los reactivos del cuestionario deben ser tan precisos que el entrevistado sepa con exactitud lo que le pregunta el investigador.

Las posibilidades de entender mal son innumerables y ningún investigador está inmune (Polivka y Rothgeb, 1993). En Estados Unidos, uno de los proyectos de investigación más reconocidos es la permanente "Encuesta de la Población Actual" de la Oficina del Censo (EPA), que mide, entre otros datos cruciales, el índice nacional de desempleo.

Una parte de la medición de la estructura del empleo se concentra en las actividades del entrevistado durante "la semana pasada", por la que la Oficina del Censo entiende del domingo al sábado. Los estudios emprendidos para determinar la exactitud de la encuesta descubrieron que más de la mitad de los interrogados tomaron "la semana pasada" como del lunes al viernes. Igualmente, mientras que la Oficina del Censo define como "trabajo de tiempo completo" a uno que ocupe 35 o más horas a la semana, los mismos estudios de evaluación mostraron que los entrevistados asumían la definición más tradicional, de 40 horas por semana. En consecuencia, la redacción de estas preguntas de la EPA se modificó en 1994 para especificar las definiciones de la Oficina del Censo.

El uso del término "nativo estadounidense" para significar "indígena estadounidense" provoca a menudo una representación extrema de este grupo étnico en las encuestas. En efecto, muchos interrogados entienden que el término significa "nacido en Estados Unidos".

### **Evite las preguntas dobles**

Frecuentemente, los investigadores les piden a los entrevistados una respuesta única para una combinación de preguntas. Esto parece ocurrir sobre todo cuando el investigador se ha identificado personalmente con una pregunta complicada. Por ejemplo, supongamos que usted pregunta a sus entrevistados si están de acuerdo o no con el enunciado "Estados Unidos deben abandonar su programa espacial y destinar el dinero a proyectos nacionales". Aunque algunos estarán

## Preguntas dobles y más allá

Incluso los investigadores profesionales prestigiados formulan a veces preguntas dobles o peores. Considere esta pregunta, que se planteó a los ciudadanos estadounidenses en abril de 1986, cuando las relaciones entre ese país y Libia estaban en un punto especialmente bajo. Algunos observadores postulaban que Estados Unidos debía entablar una guerra abierta con la pequeña nación del norte de África. Un sondeo de Harris trató de averiguar la opinión del público.

Si Libia incrementa ahora sus actos terroristas contra Estados Unidos y nosotros le infligimos más daño a Libia, inevitablemente acabaremos por entrar en guerra e invadir aquel país, lo que estaría mal.

Los entrevistados tenían la posibilidad de responder “de acuerdo”, “en desacuerdo” y “no

estoy seguro”. Observe los elementos comprendidos en el complicado enunciado:

1. ¿Incrementará Libia sus actos terroristas contra Estados Unidos?
2. ¿Infligirá más daño Estados Unidos a Libia?
3. ¿Es inevitable o no que Estados Unidos entre en guerra?
4. ¿Invadiría Estados Unidos a Libia?
5. ¿Estaría bien o mal?

Todos estos elementos abren la posibilidad de muchos puntos de vista, muchos más que las tres alternativas que la encuesta ofrecía a los entrevistados. Aunque supusiéramos como hipótesis que Libia incrementara “sus actos terroristas” y que como respuesta Estados Unidos infligiera “más daño a Libia”, uno tendría cualquiera de las siguientes siete expectativas sobre el resultado:

inequívocamente de acuerdo y otros no, algunos más no sabrán responder. Unos quisieran que se abandonara el programa espacial y se devolviera el dinero a los contribuyentes, en tanto que otros preferirían que continuara el programa pero que también se gastara dinero en proyectos nacionales. Todos estos últimos no estarían de acuerdo ni en desacuerdo sin confundirlo.

Como regla general, cada vez que aparezca la palabra y en una pregunta o enunciado del cuestionario, debe verificar si acaso está formulando una *pregunta doble*. Véanse algunas variaciones imaginativas sobre el tema en el recuadro “Preguntas dobles y más allá”.

### **Los interrogados deben ser capaces de responder**

Al pedirle a los entrevistados que le den información, debe preguntarse constantemente

si son capaces de ofrecérsela en forma confiable. En un estudio de la educación infantil, podría preguntarles a qué edad empezaron a hablar con sus padres. Aparte del problema de definir *hablar con sus padres*, es dudoso que la mayoría de los entrevistados lo recuerde con algún grado de precisión.

Tomemos otro ejemplo. Los líderes de los grupos estudiantiles a veces piden a sus electores que indiquen la manera en que deben gastarse sus cuotas. Habitualmente, se les pide que indiquen el porcentaje de los fondos disponibles que hay que asignar a una larga lista de actividades. Sin un conocimiento bastante bueno de la naturaleza de tales actividades y sus costos, los entrevistados no pueden dar respuestas significativas (los *gastos administrativos* recibirán poco apoyo, aunque sean esenciales para el conjunto del programa).

	E.U. no entrará en guerra	La guerra es probable	La guerra es inevitable
E.U. no invadirá			
Libia	1	2	3
E.U. invadirá			
Libia pero estaría mal		4	5
E.U. invadirá			
Libia pero estaría bien		6	7

Este examen de los pronósticos sobre la situación libia no es el único ejemplo de preguntas dobles que se cuelean en las investigaciones de la opinión pública. Las siguientes son algunas preguntas que el sondeo de Harris formuló para tratar de medir la opinión pública estadounidense sobre el entonces

secretario general de la Unión Soviética, Gorbachov:

Parece el tipo de gobernante ruso que acepta que tanto los soviéticos como los estadounidenses pueden destruirse unos a otros con los misiles nucleares, por lo que es mejor llegar a acuerdos verificables de control de armas.

Parece moderno, educado y atractivo, lo que es un buen signo para la paz en el mundo.

Aunque se ve mucho más moderno y atractivo, sería un error pensar que será muy diferente de otros gobernantes rusos.

¿Cuántos elementos encuentra usted en cada enunciado? ¿Cuántas opiniones podría tener la gente en cada caso? ¿Qué significan los simples "de acuerdo" y "en desacuerdo" en estos casos?

Fuente: Informado en *World Opinion Update*, octubre de 1985 mayo de 1986, respectivamente.

Un grupo de investigadores que examinaba la experiencia de los adolescentes para manejar insistía en formular una pregunta abierta sobre los kilómetros que habían recorrido desde que obtuvieron su licencia. Algunos consejeros argumentaron que pocos conductores serían capaces de calcular con exactitud tal información, pero se planteó la pregunta de todas maneras. Así, algunos adolescentes dijeron que habían manejado cientos de miles de kilómetros.

### **Los interrogados deben estar dispuestos a responder**

A menudo nos gustaría saber asuntos de personas que no están dispuestas a compartírnoslos con nosotros. Por ejemplo, Yajie Bian indica que muchas veces se le ha dificultado obtener respuestas francas de los chinos,

pues la gente acostumbra tener cuidado con lo que dice en situaciones públicas con el fin de sobrevivir en un régimen autoritario. Por ejemplo, entre 1966 y 1976, durante la Revolución cultural, a causa del radicalismo político y la intensidad política de todo el país, era casi imposible emplear las técnicas de encuestas para reunir datos válidos y confiables dentro de China sobre las experiencias, características y actitudes del pueblo sobre el régimen comunista.

(1994: 19-20)

En ocasiones, los estadounidenses interrogados dicen que no se han decidido cuando, de hecho, ya tienen una opinión, pero piensan que son la minoría. En tales condiciones, son reacios a confiarle su opinión a un extraño (el entrevistador). Debido a este problema, la Organización Gallup, por ejemplo, ha empleado un formato de "boleta secreta", que simula unas elecciones reales en las que el "votante" goza de un anonimato completo. En un análisis de los datos de los sondeos

## 130 Capítulo 6 Operacionalización

electorales de Gallup entre 1944 y 1988. Andrew Smith y G. F. Bishop (1992) descubrieron que esta técnica reducía sustancialmente el porcentaje de sujetos que se decían indecisos sobre el candidato por quien votarían.

Este problema no se limita a las encuestas. Richard Mitchell (1991: 100) enfrentó un problema parecido en su investigación de campo de los supervivientes estadounidenses.

Por ejemplo, los supervivientes se muestran ambivalentes sobre revelar sus identidades e inclinaciones. Se dan cuenta de que el anonimato los protege del ridículo ante una mayoría incrédula, pero el aislamiento obligado reduce sus oportunidades de reclutamiento e intercambio de información..

Los supervivientes "secretos" evitan los teléfonos, "lavan" su correspondencia mediante intercambio de cartas, usan sobrenombres y alias y esconden cuidadosamente sus direcciones a los desconocidos. Sin embargo, una ocasión en que fui invitado a las reuniones del grupo, resultaron ser cooperadores en las entrevistas.

### ***Las preguntas deben ser pertinentes***

Del mismo modo, las preguntas del cuestionario deben ser pertinentes para la mayoría de los entrevistados. Cuando se piden actitudes acerca de un tema en el que pocos interrogados han pensado o que en realidad les importa poco, no es probable que los resultados sean muy útiles. Desde luego, como los entrevistados pueden expresar actitudes aunque nunca le hayan dedicado un pensamiento al tema, se corre el riesgo de equivocarse.

Este punto se ejemplifica cuando los investigadores formulan preguntas con respuestas que se refieren a personajes y temas ficticios. En un sondeo político que realicé, le pregunté a los entrevistados si conocían a 15 figuras políticas de la comunidad. Como ejercicio metodológico, inventé un nombre: Tom Sakumoto. En respuesta, nueve por ciento de los interrogados dijo que lo conocía. De éstos, aproximadamente la mitad dijo que lo había visto en la televisión o que había leído de él en el periódico.

Si usted obtiene respuestas de temas ficticios, puede descartarlas. Pero cuando el tema es real, quizá no tenga forma de saber cuáles son las

respuestas que reflejan genuinamente las actitudes y cuáles son respuestas sin sentido a preguntas irrelevantes. En términos ideales, quisiéramos que los interrogados dijeran que no saben, no tienen opinión o están indecisos cuando así fuera el caso; sin embargo, como hemos visto, a menudo inventan las respuestas.

### ***Los reactivos cortos son mejores***

Con objeto de ser preciso e inequívoco y resaltar la importancia de un tema, el investigador se enreda a veces en reactivos largos y complicados. Hay que evitarlo. Los entrevistados no están dispuestos a estudiar un reactivo para comprenderlo. Deben ser capaces de leerlo rápidamente, entender su sentido y elegir o dar una respuesta sin dificultades. En general, debemos suponer que los interrogados leerán los reactivos y darán las respuestas con rapidez; por tanto, hay que ofrecer reactivos cortos y claros que no se interpreten mal en esas condiciones.

### ***Evite los reactivos negativos***

La aparición de una negación en un reactivo facilita los errores de interpretación. Al pedir su acuerdo o desacuerdo con el enunciado "Estados Unidos *no* debe reconocer a Cuba", una parte considerable de los entrevistados saltará la palabra *no* y responderá en consecuencia. Así, algunos estarán de acuerdo con el enunciado si están en favor del reconocimiento y otros también lo aceptarán porque se oponen. Y uno nunca sabrá cuál es cuál.

En un estudio del apoyo a las libertades civiles, se preguntó a los entrevistados si pensaban que "se debía prohibir a las siguientes personas que enseñaran en las escuelas públicas", y se presentaba una lista que incluía comunistas, miembros del Ku Klux Klan, etc. Al lado de cada entrada se daban las categorías de respuesta "sí" o "no". Una comparación de las respuestas de este reactivo con las de otros que reflejaban apoyo a las libertades civiles sugería con fuerza que muchos entrevistados dieron la respuesta "sí" para indicar su disposición a que tal persona enseñara, más que señalar que habría que prohibirle la docencia (un estudio posterior de la serie que proporcionaba como respuestas las categorías "permitir" y "prohibir" produjo resultados mucho más claros).

**Evite reactivos y términos tendenciosos**

Recuerde, de nuestro análisis anterior de la conceptualización y la operacionalización, que en última instancia no hay significados verdaderos para ninguno de los conceptos que estudiamos en las ciencias sociales. *Prejuicios* no tiene una definición última correcta, y el hecho de que una persona sea prejuiciosa o no depende de nuestra definición del término. El mismo principio general se aplica a las respuestas que tenemos de quienes contestan los cuestionarios.

El significado de la respuesta a una pregunta depende en buena parte de su redacción. Esto es cierto para toda pregunta y respuesta. Algunas preguntas parecen fomentar ciertas respuestas más que otras. Las preguntas que estimulan a los interrogados a responder de cierta manera son *tendenciosas*.

Muchos investigadores aceptan el efecto probable de una pregunta que comienza "¿no está de acuerdo con el presidente en que...?", y ninguno que se respete emplearía tal reactivo. Por desgracia, el efecto tendencioso de reactivos y términos es mucho más sutil que en este ejemplo.

La mera identificación de una actitud o posición con una persona o institución de prestigio puede volver tendenciosas las preguntas. El reactivo "¿está de acuerdo o no con la decisión reciente de la Suprema Corte de que...?" tendría un efecto similar. Ahora bien, no quiero decir que esta redacción producirá necesariamente un consenso o siquiera una mayoría que apoye la posición identificada con la persona o institución prestigiada, sino sólo que es probable que tal apoyo sea mayor que el que se obtendría sin esa identificación.

A veces, el efecto de las diferentes formas de redactar una pregunta es relativamente sutil. Por ejemplo, cuando Kenneth Rasinski (1989) analizó los resultados de varios estudios de la Encuesta Social General sobre las actitudes hacia los gastos del gobierno, descubrió que la manera de identificar los programas afectaba el grado de apoyo público que recibían. Veamos algunas comparaciones:

Más apoyo	Menos apoyo
"Ayuda a los pobres"	"Asistencia social"
"Detener el aumento de los índices de delincuencia"	"Aplicación de la ley"
"Enfrentar la drogadicción"	"Rehabilitación de drogadictos"
"Resolver los problemas de las grandes ciudades"	"Auxiliar a las ciudades grandes"
"Mejorar las condiciones de los negros"	"Asistencia a los negros"
"Defender la seguridad social"	"Seguridad social"

Por ejemplo, en 1986, 62.8 por ciento de los entrevistados dijo que se gastaba muy poco en "ayuda a los pobres", mientras que en una encuesta equivalente ese mismo año, sólo 23.1 por ciento dijo que se gastaba muy poco en "asistencia social".

En este contexto, por lo general usted debe estar atento a lo que los investigadores llaman la *deseabilidad social* de las preguntas y las respuestas. Cuando pide información a los demás, éstos responden a través de un filtro que los hará verse mejor. Esto es cierto sobre todo en las encuestas en persona. Así, por ejemplo, cierto hombre podría pensar que las cosas estarían mucho mejor si las mujeres se quedaran en la cocina, no tuvieran derecho al voto, estuvieran obligadas a guardar silencio en público, etc. Sin embargo, al ser interrogado sobre su apoyo a la igualdad de derechos para las mujeres, no le gustaría aparecer como un cerdo macho chovinista. Como reconocería que sus opiniones habrían sido progresistas en el siglo XV pero que están fuera de tono con las ideas actuales, escogería decir "sí". El mejor lineamiento que puedo ofrecerle en relación con este tema es sugerirle que se imagine cómo se sentiría usted de -responder a las preguntas que les plantea a sus entrevistados. Si se siente avergonzado, pervertido, inhumano, estúpido, irresponsable o algo por el estilo, debe meditar con seriedad si ellos estarán dispuestos a dar tales respuestas.

Suele ser difícil anticipar el efecto tendencioso de ciertas redacciones. En encuestas y experimentos, a veces es útil pedir a los interrogados que consideren situaciones hipotéticas y respondan cómo creen que actuarían. Como tales situaciones atañen a otras personas, los nombres utilizados pueden influir en las respuestas. Por ejemplo, los investigadores saben desde hace mucho que los nombres masculinos para las situaciones hipotéticas producen respuestas diferentes que los femeninos. Las investigaciones de Joseph Kasof (1993) destacan la importancia que tienen determinados nombres. Si en general suscitan imágenes positivas o negativas en cuanto a atractivo, edad, inteligencia, etc. La revisión de Kasof de investigaciones anteriores indica que ha habido una tendencia a usar nombres más valorados para los hombres que para las mujeres.

Como en todos los demás ejemplos, debe examinar cuidadosamente el propósito de su investigación. No cometa nunca el error de pensar que hay en definitiva formas "correctas" e "incorrectas" de formular las preguntas. Más aún, cuando tenga

dudas sobre la mejor manera de plantear una pregunta, recuerde que debe formular más de una.

En fin, estos son algunos lineamientos generales para redactar preguntas que arrojen datos para su análisis e interpretación. En seguida nos ocuparemos de la elaboración de cuestionarios (en el capítulo 10 hay más información sobre la aplicación de cuestionarios).

## Elaboración de cuestionarios

Así como hemos visto que hay muchas formas de operacionalizar las variables de la investigación social, los cuestionarios se usan en conexión con muchos modos de observación. Los cuestionarios son esenciales para las encuestas y se vinculan más a éstas, pero también se aplican mucho en experimentos, investigaciones de campo y otras actividades de recopilación de datos.

Debido al uso difundido de los cuestionarios en la investigación social, dirigiremos nuestra atención a algunas técnicas establecidas para elaborarlos. Debe leer las secciones siguientes no sólo como una continuación de nuestros análisis teóricos de la conceptualización y la medición, sino también con el fin de aprender una destreza práctica.

Al igual que en nuestro estudio anterior sobre la redacción de preguntas, voy a correr el riesgo de ofenderlo presentándole algunos detalles esenciales que acaso le parezcan indignos de atención científica, así como otros que le resulten demasiado obvios para mencionarlos. Corro este riesgo porque yo he cometido todos los errores de los que quiero precaverlo y he visto a los demás hacer otro tanto. Comencemos con algunos puntos sobre los formatos de los cuestionarios.

### **Formato general de los cuestionarios**

El formato de un cuestionario es tan importante como la naturaleza y la redacción de las preguntas. Un cuestionario mal presentado puede hacer que los interrogados se salten preguntas, se confundan sobre la naturaleza de los datos deseados e incluso lo desechen.

Como regla general, el cuestionario no debe estar apiñado, sino disperso. Los investigadores inexpertos suelen temer que su cuestionario parezca demasiado largo; por eso, comprimen

varias preguntas en una sola línea, abrevian otras y tratan de usar tan pocas páginas como sea posible. Todos estos esfuerzos tienen poco tino e incluso son peligrosos. Colocar más de una pregunta en una sola línea hará que algunos entrevistados se salten por completo la segunda. Algunos interpretarán mal las preguntas abreviadas. Más en general, los entrevistados que sientan que perdieron mucho tiempo en la primera página de lo que parecía un cuestionario breve estarán más desanimados que quienes terminen pronto las primeras páginas de un cuestionario que al principio parecía largo. Además, estos últimos cometerán menos errores y no estarán obligados a leer de nuevo las preguntas abreviadas confusas ni a escribir una respuesta grande en un espacio pequeño.

No es posible exagerar la importancia de extender el espacio de las preguntas del cuestionario. Comprimir los cuestionarios es desastroso, sea que los deban llenar los propios entrevistados o bien que los apliquen entrevistadores capacitados. Además, el procesamiento de estos cuestionarios es otra pesadilla. Diremos más del tema en el capítulo 14.

### **Formato para los entrevistados**

En uno de los formatos de cuestionario más comunes, se espera que el entrevistado marque una respuesta de una serie. De los diversos métodos disponibles, es mi experiencia que los recuadros bien espaciados son los mejores. Los procesadores de texto modernos hacen de los recuadros una técnica práctica en nuestros días; para los tipos de los recuadros también es fácil y limpio. Incluso en una máquina de escribir se pueden trazar recuadros aproximados.

Si el cuestionario se escribe en una máquina mecanográfica con corchetes, se logran recuadros excelentes con un corchete de apertura, un espacio y uno de cierre: [ ]. Si no cuenta con corchetes, los paréntesis funcionarán razonablemente bien de la misma manera: ( ). En cambio, no aconsejo el uso de diagonales y rayas. Primero, esta técnica exige mucho más esfuerzos de mecanografía, y segundo, el resultado no es muy limpio, en particular si las categorías de respuestas deben estar separadas por un solo espacio. En la figura 6.2 se muestra una comparación de varios métodos.

De los tres métodos que se muestran, los corchetes y los paréntesis se ven sin duda más limpios; el método de las diagonales y las rayas se ve desaliñado. Como todas las máquinas de escribir

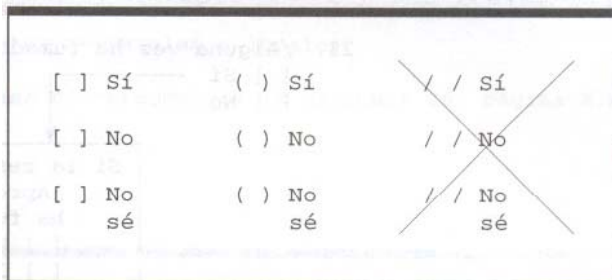
tienen por lo menos paréntesis, no hay excusa para emplear diagonales y rayas. El peor método es dejar espacios en blanco para marcar, porque los entrevistados pondrán marcas más bien grandes y no será posible determinar cuál es su respuesta.

Si prepara un cuestionario en una computadora, dedique unos minutos extra a crear recuadros genuinos que le den a su trabajo una apariencia más profesional. Éstos son unos ejemplos sencillos:

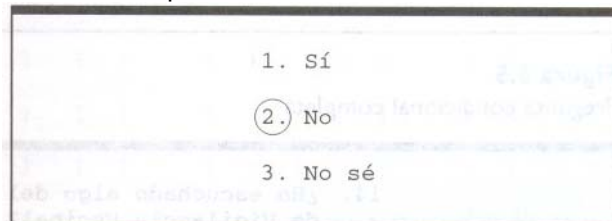


En lugar de ofrecer recuadros para marcar, el investigador puede imprimir números de código junto a cada respuesta y pedir a los interrogados que circulen el apropiado (véase la figura 6.3). Este método tiene la ventaja adicional de especificar un número de código para manejarlo después, en la etapa de procesamiento (véase el capítulo 14). Ahora bien, si se van a circular los números, se deben dar instrucciones claras y visibles, porque muchos entrevistados tacharán el número apropiado, lo que dificulta aún más el procesamiento de datos (observe que la técnica se emplea con más seguridad si los entrevistadores aplican el cuestionario, dado que pueden instruir y supervisar a los interrogados).

**Figura 6.2**  
Tres formatos de respuesta



**Figura 6.3**  
Circular la respuesta



## Preguntas condicionales

Con mucha frecuencia, ciertas preguntas de los cuestionarios serán pertinentes para algunos de los entrevistados e irrelevantes para otros. Por ejemplo, en un estudio de los métodos de control natal no queremos preguntarle a los hombres si toman píldoras anticonceptivas.

Se da esta situación cuando uno quiere formular una serie de preguntas sobre cierto tema. Digamos que usted quiere preguntar a sus entrevistados si pertenecen a determinada organización y, si es el caso, con qué frecuencia asisten a las reuniones, si han tenido algún puesto directivo, etc. O tal vez quiera saber si han oído de cierto asunto político y cuáles son las opiniones de quienes lo han hecho.

Las preguntas sucesivas de estas series se llaman **preguntas condicionales**: si deben plantearse y responderse depende de la respuesta a la primera pregunta de la serie. El uso apropiado de las preguntas condicionales facilita la tarea de los entrevistados de contestar el cuestionario, pues no están obligados a tratar de responder preguntas que son irrelevantes para ellos.

Hay varios formatos para las preguntas condicionales. El que se muestra en la figura 6.4 es el más claro y eficaz. Observe que tiene dos elementos claves. Primero, la pregunta condicionada queda aislada de las otras, desplazada a un lado y encerrada en un recuadro. Segundo, una flecha conecta la pregunta condicionada con la respuesta de que depende. En la ilustración, se pretende que sólo los entrevistados que respondan *si* pasen a responder la pregunta condicionada. El resto simplemente se la salta.

Advierta que las preguntas de la figura 6.4 podrían haberse planteado en una sola pregunta: "¿Cuántas veces, si acaso, ha fumado marihuana?". Así, las categorías de respuesta habrían sido "nunca", "una vez", "2 a 5 veces", etc. Esta pregunta única se aplicaría a todos los entrevistados y cada uno escogería la categoría de respuesta apropiada. Sin embargo, tal vez la pregunta los presione de alguna forma a informar que han fumado marihuana, puesto que se quiere saber cuántas veces la han fumado, aunque permita la posibilidad de *casos excepcionales que nunca han fumado marihuana ni siquiera una vez* (las cursivas de la oración anterior dan una mediana indicación de la forma en que los entrevistados leerían la pregunta). El formato de



Figura 6.4

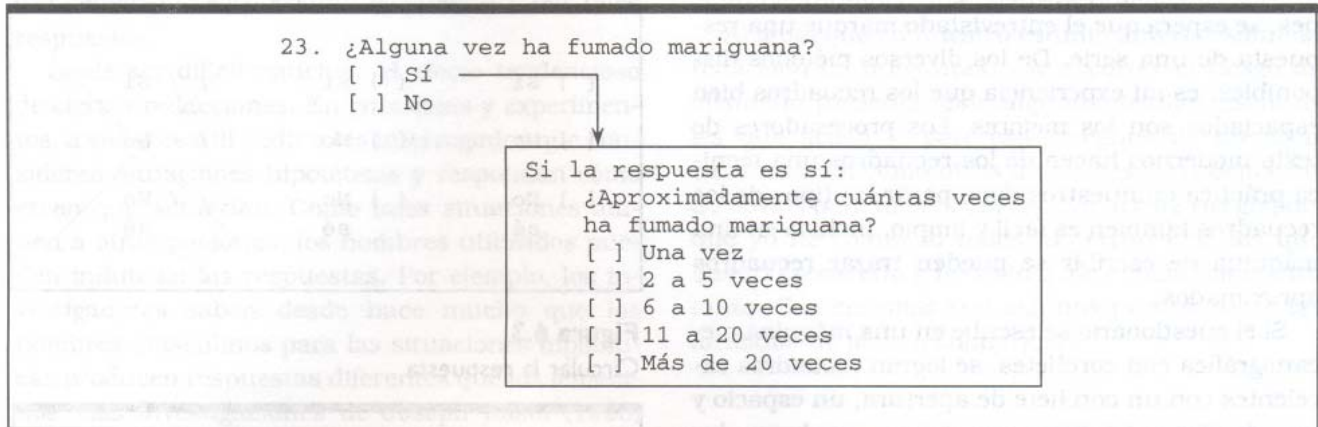
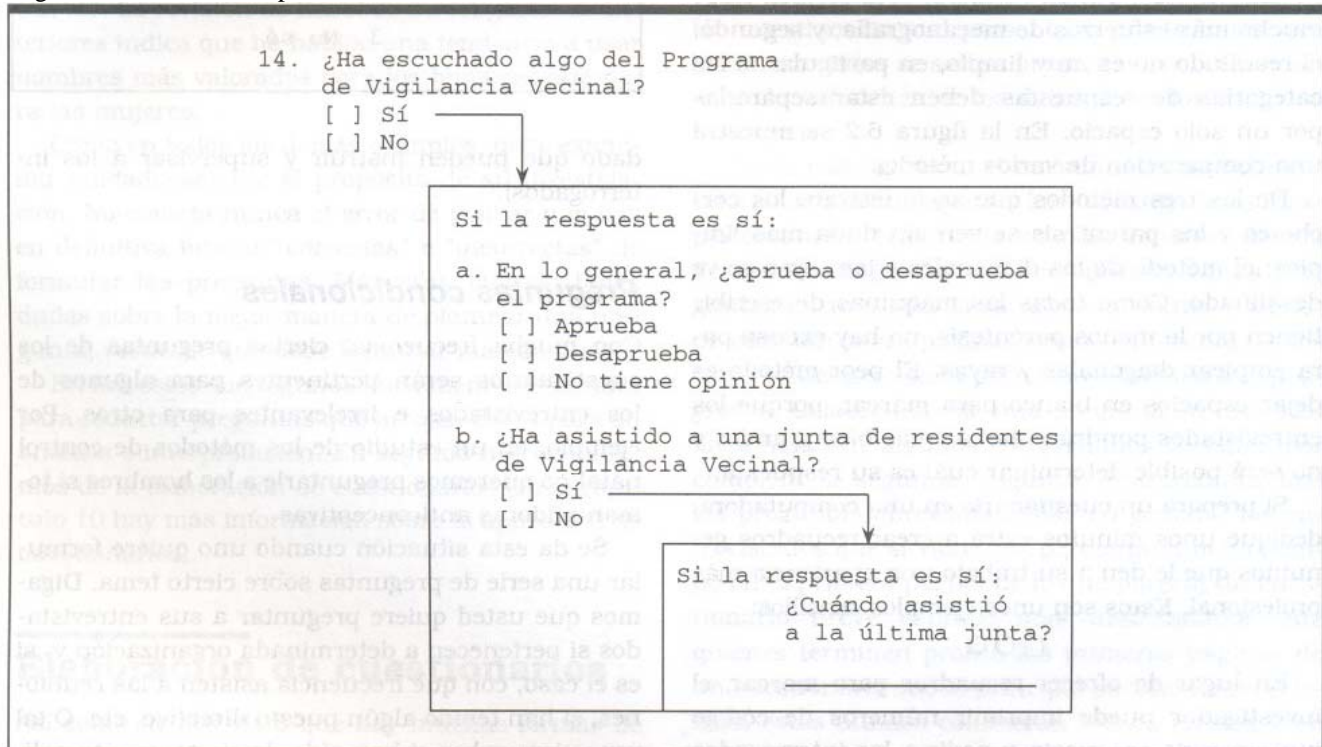


Figura 6.5

Pregunta condicional completa



pregunta condicionada de la figura 6.4 debe reducir la sutil presión sobre los entrevistados para que digan que fumaron marihuana. Los comentarios precedentes deben mostrar que los temas aparentemente teóricos de *validez* y *confiabilidad* tienen que ver con la materia tan mundana de cómo poner unas preguntas en una hoja de papel.

Usadas apropiadamente, incluso las series complicadas de preguntas condicionales pueden

organizarse sin confundir al entrevistado. La figura 6.5 ilustra un ejemplo más complejo.

A veces, una serie de preguntas condicionales es tan grande que se extiende varias páginas. Supongamos que usted estudia las actividades políticas de los estudiantes universitarios y quiere formular muchas preguntas a quienes han votado en unas elecciones nacionales, estatales o locales. Podría separar a los entrevistados pertinentes con una pregunta inicial, como "¿ha votado en algunas

**Figura 6.6**

Instrucciones para saltar

13. ¿Ha votado en algunas elecciones nacionales, estatales o locales?  
 Sí (responda las preguntas 14 a 25)  
 No (Salte las preguntas 14 a 25. Pase directamente a la pregunta 26, página 8.)

**Figura 6.7**

Formato de matriz de preguntas

17. Junto a cada uno de los enunciados siguientes, indique si está Muy de acuerdo (MA), De acuerdo (A), En desacuerdo (D), Muy en desacuerdo (MD) o Indeciso (I).

	<u>MA</u>	<u>A</u>	<u>D</u>	<u>MD</u>	<u>I</u>
a. Lo que este país necesita . . . es más ley y orden . . . . .	[ ]	[ ]	[ ]	[ ]	[ ]
b. Hay que desarmar a la policía del país . . . . .	[ ]	[ ]	[ ]	[ ]	[ ]
c. Durante las revueltas, hay que disparar a los saqueadores . . etc.	[ ]	[ ]	[ ]	[ ]	[ ]

elecciones nacionales, estatales o regionales?", pero sería confuso situar las preguntas *condicionales* en un recuadro que se prolongara varias páginas. Sería más sensato anotar después de cada respuesta las instrucciones entre paréntesis para indicar a los entrevistados que contesten o salten las preguntas condicionales. La figura 6.6 ilustra este método.

Además de estas instrucciones, valdría la pena situar otras en la parte superior de cada página que contuviera sólo las preguntas condicionales. Por ejemplo, se diría: "Esta página es sólo para los que han votado en unas elecciones nacionales, estatales o locales". Las instrucciones claras ahorran a los entrevistados la frustración de leer y devanarse los sesos con preguntas que son irrelevantes para ellos, además de que aumenta la probabilidad de que respondan aquellos para quienes son pertinentes.

## **Matriz de preguntas**

Muy a menudo querrá formular preguntas que tengan el mismo conjunto de categorías de respuesta. Ésta suele ser la situación cuando se aplican las categorías de Likert. En tales

casos, es posible elaborar una matriz de reactivos y respuestas, como se ilustra en la figura 6.7.

Este formato ofrece varias ventajas sobre los demás. Primero, aprovecha eficazmente el espacio. Segundo, los entrevistados responderán más rápido una serie de preguntas presentadas de esta manera. Además, el formato aumenta la capacidad tanto de los interrogados como del entrevistador de comparar las respuestas que se dieron a las preguntas. Como los entrevistados pueden repasar en forma rápida sus respuestas a los reactivos anteriores de la serie, están en posición de escoger, digamos, entre "muy de acuerdo" y "de acuerdo" para responder determinado enunciado comparando el grado de su asentimiento en las respuestas que ya dieron.

Sin embargo, hay algunos peligros inherentes en el uso de este formato. Sus ventajas pueden animarlo a estructurar un reactivo de modo que sus respuestas se ajusten al formato matricial cuando sería mejor un conjunto de respuestas diferente y más propio. Asimismo, el formato matricial puede fomentar entre algunos interrogados *una predisposición de respuesta*: adoptan la costumbre de, digamos,

estar de acuerdo con todos los enunciados.

Esto es especialmente probable si la serie de enunciados comienza con varios que indican cierta tendencia (por ejemplo, un punto de vista político liberal) y sólo unos cuantos al último representan la postura opuesta. Los entrevistados suponen que todos los enunciados tienen la misma tendencia y, al leer rápido, se equivocan en algunos y dan respuestas erróneas. Para reducir el problema en alguna medida, se alternan enunciados que representen orientaciones distintas y se redactan de manera clara y breve.

### ***El orden de los reactivos del cuestionario***

El orden en que se presentan los reactivos del cuestionario también influye en las respuestas. Primero, la apariencia de una pregunta tiene un efecto en las respuestas de las siguientes. Por ejemplo, si se han formulado varias preguntas sobre el peligro del comunismo en Occidente y luego se pide a los interrogados que indiquen (pregunta abierta) aquello que en su opinión plantea un peligro para Occidente, se citará más el comunismo que en otras ocasiones. En esta situación, es preferible colocar primero la pregunta abierta.

Si se pide a los entrevistados que evalúen su religiosidad en general ("¿cuán importante es para usted la religión en general?"), sus respuestas a preguntas posteriores que conciernen a aspectos concretos de la religiosidad aspirarán a ser congruentes con la evaluación del principio. Lo contrario también es cierto. Si se les formulan primero preguntas sobre varios aspectos de su religiosidad, una evaluación general posterior reflejaría aquellas respuestas.

El efecto del orden de los reactivos no es uniforme. Cuando J. Edwin Benton y John Daly (1991) realizaron una encuesta del gobierno local, descubrieron que los entrevistados con menos educación resentían más la influencia del orden de los reactivos que quienes tenían más.

Algunos investigadores tratan de superar este efecto colocando los reactivos al azar, pero suele ser un esfuerzo vano. Para empezar, un conjunto aleatorio de reactivos le parecerá caótico e inútil a los interrogados. Además, será difícil de contestar, puesto que deben cambiar constantemente su atención de un tema a otro. Por último, incluso unos reactivos en orden aleatorio tendrán el efecto que hemos señalado, con la excepción de que usted no tendrá el control sobre tal efecto.

La solución más segura es la sensibilidad ante el problema. Aunque no pueda evitar el efecto del orden de los reactivos, trate de calcularlo. Así, estará en posición de darle a los resultados una interpretación significativa. Si el orden de los reactivos parece una cuestión importante en cierto estudio, elabore varias versiones del cuestionario con diferentes ordenamientos. Entonces, podrá determinar los efectos. Al menos, debe realizar una prueba previa del cuestionario en las diversas formas.

El orden conveniente de los reactivos difiere un tanto entre las entrevistas y los cuestionarios de aplicación personal. En éstos, lo mejor suele ser comenzar con el conjunto más interesante de reactivos. Los posibles interrogados que echen una ojeada casual a los primeros reactivos deben *querer* contestarlos. Quizá los reactivos piden opiniones que ansían expresar. Sin embargo, al mismo tiempo no deben ser reactivos amenazadores (sería una mala idea comenzar con reactivos sobre la conducta sexual o el consumo de drogas). Los datos demográficos aburridos (edad, sexo, etc.) deben colocarse al final del cuestionario de aplicación personal. Si se ponen al principio, como se sienten tentados a hacer muchos investigadores inexpertos, el cuestionario adquiere la apariencia inicial de una forma rutinaria, y quien lo recibe quizá no se sienta motivado a contestarlo.

En las entrevistas, lo opuesto suele ser lo cierto. Cuando el posible entrevistado abre la puerta por primera vez, el entrevistador debe ganarse pronto su simpatía. Después de una breve introducción al estudio, lo mejor es que

comience por enumerar a los miembros de la familia y obtener los datos demográficos de cada uno. Estos reactivos se responden con facilidad y en general no son amenazadores. Cuando se ha logrado la compenetración inicial, el entrevistador puede pasar a la zona de actitudes y temas más delicados. El entrevistador que comience con la pregunta "¿cree usted en la brujería?" terminará más bien rápido.

## **Instrucciones**

Todo cuestionario, sea que lo respondan los interrogados o lo apliquen los entrevistadores, debe contener instrucciones claras y comentarios introductorios, si es conveniente.

Es útil comenzar los cuestionarios de aplicación personal con las instrucciones básicas para llenarlo. Aunque en estos días muchas personas están bastante familiarizadas con formas y cuestionarios, hay que comenzar diciéndoles exactamente lo que se quiere: que para indicar sus respuestas a ciertas preguntas marquen con una paloma o una X el recuadro junto a la respuesta adecuada, o que la escriban cuando así se pida. Si hay muchas preguntas abiertas, los entrevistados deben tener algunos lineamientos sobre la extensión de las respuestas. Si usted quiere animar a sus entrevistados a que elaboren sus repuestas a preguntas abiertas, hágalo notar.

Si el contenido de un cuestionario está organizado en secciones, tendencias políticas, opiniones religiosas, antecedentes, presente cada sección con una breve exposición de su contenido y propósitos. Por ejemplo, "en esta sección nos gustaría saber cuáles son los problemas que los vecinos consideran más importantes de la comunidad". Los reactivos demográficos al final del cuestionario de aplicación personal pueden presentarse así: "Por último, quisiéramos saber un poco sobre usted para que conozcamos las opiniones de distintas personas acerca de los temas que examinamos".

Las introducciones breves ayudan al entrevistado a comprender el cuestionario. Hacen que parezca menos caótico, en particular si busca una variedad de datos y le permiten al entrevistado adoptar la actitud intelectual adecuada para contestar las preguntas.

Algunas preguntas requieren instrucciones especiales para facilitar las respuestas correctas. Esto es así sobre todo si determinada pregunta difiere de las instrucciones generales que atañen al conjunto del cuestionario. Algunos ejemplos concretos ilustrarán la situación.

A pesar de la conveniencia de las categorías de respuestas mutuamente excluyentes en las preguntas abiertas, ocurrirá que más de una respuesta se aplique a los entrevistados. Si usted quiere una *única* respuesta, debe quedar perfectamente claro en la pregunta; por ejemplo: "En la lista que sigue, marque su *principal* razón para asistir a la universidad". Con frecuencia, la pregunta principal puede estar seguida de una nota entre paréntesis: "(marque la mejor respuesta)". Por lo demás, si quiere que los entrevistados marquen tantas respuestas como convenga, también debe dejarlo claro.

Cuando el entrevistado deba ordenar un conjunto de categorías de respuesta, las instrucciones deben indicarlo así y debe emplearse otro formato de respuesta (por ejemplo, espacios en blanco y no recuadros). Estas instrucciones indicarán la manera en que se deben clasificar muchas respuestas (digamos, todas, la primera y la segunda, la primera y la última, la más y la menos importante) y el orden de la clasificación (por ejemplo, "escriba un 1 junto a la más importante, un 2 junto a la siguiente en importancia, etc."). Ahora bien, el ordenamiento de las respuestas suele ser difícil para los entrevistados, porque tienen que leer la lista una y otra vez, así que esta técnica debe dejarse sólo para aquellas situaciones en las que ningún otro método producirá el resultado deseado.

## 138 Capítulo 6 Operacionalización

En las matrices de preguntas que constan de varias partes, es útil dar instrucciones especiales a menos que se use el mismo formato en todo el cuestionario. A veces se desea que los entrevistados marquen una respuesta en cada *columna* de la matriz; en otros cuestionarios se pide que marquen una en cada *fila*. Si el cuestionario contiene ambos tipos, es útil añadir las instrucciones que aclaren lo que se espera en cada caso.

### **Prueba previa del cuestionario**

No importa el cuidado con que diseñe usted un instrumento de recopilación de datos, como un cuestionario, siempre cabe la posibilidad de hecho, la certeza de cometer errores. Es seguro que incurrirá en equivocaciones: una pregunta ambigua, una que no se puede contestar o alguna otra infracción a las reglas que hemos estudiado.

La mejor protección contra estos errores es realizar una *prueba previa* de todo el cuestionario, o de una parte de él. Déle el cuestionario a 10 personas de, por ejemplo, su torneo de boliche. No es esencial que los sujetos de la prueba previa comprendan una muestra representativa, pero debe acudir a personas para quienes el cuestionario sea por lo menos pertinente.

En general, lo mejor es que les pida a las personas que contesten el cuestionario y no que lo lean en busca de errores. Ocurre con mucha frecuencia que una pregunta parece tener sentido en la primera inspección, pero luego se descubre que no es posible contestada con exactitud.

Stanley Presser y Johnny Blair (1994) exponen varias estrategias de pruebas previas y señalan la eficacia de cada una. También ofrecen datos sobre los costos de cada método.

Si le diera todas las sugerencias y lineamientos que tengo para la elaboración de un cuestionario, esta sección sería más larga que el resto del libro. Así, un tanto a regañadientes, termino este estudio con un ejemplo de un cuestionario real en el que se muestra la utilidad de estos comentarios en la práctica.

Sin embargo, antes de pasar a la ilustración, quiero mencionar un aspecto crucial del diseño de cuestionarios que estudiaremos en el capítulo 14: la *precodificación*. Como la información que se reúne en los cuestionarios se vierte en alguna clase de formato de computadora, suele ser apropiado

incluir en el propio cuestionario las instrucciones del procesamiento de los datos. Las instrucciones señalan el lugar en donde se almacenarán las unidades de información en los archivos de máquina. En el capítulo 14 veremos el carácter de este almacenamiento e indicaremos las anotaciones adecuadas. Como adelanto, observe que el siguiente ejemplo está precodificado con los misteriosos números que aparecen junto a las preguntas y las categorías de respuesta.

### **Ejemplo compuesto**

La figura 6.8 forma parte de un cuestionario utilizado por el Centro Nacional de Investigación de la Opinión de la Universidad de Chicago en su muy empleada Encuesta Social General. El cuestionario aborda las opiniones sobre el gobierno y está diseñado para la aplicación personal.

---

## **La operacionalización no cesa**

Aunque hemos estudiado la conceptualización y la operacionalización como actividades que preceden a la recopilación y el análisis de los datos, es decir, que hay que diseñar los reactivos del cuestionario antes de enviarlo, debe advertir que ambos procesos continúan durante todo el proyecto de investigación, aun cuando ya se hayan reunido los datos en una encuesta general estructurada. Como vimos, en los métodos menos estructurados, como la investigación de campo, la identificación y la especificación de los conceptos pertinentes son inseparables del proceso continuo de observación.

En el diseño de estudios que requieren mediciones estandarizadas de los conceptos, lo más sensato es aprovechar varios indicadores de cada uno, particularmente si están muy abiertos a diferentes interpretaciones y definiciones. Al medir la variable de maneras distintas, estará en posición de examinar otras definiciones operacionales durante su análisis. Como podrá escoger de varios indicadores y tendrá muchas formas de elaborar mediciones combinadas, será capaz de experimentar con diversas medidas de las que cada una represente una conceptualización y operacionalización un tanto distinta para decidir cuál da las respuestas más claras y útiles a sus preguntas.

**Figura 6.8**  
**Cuestionario de muestra**

10. Las siguientes son acciones que podría emprender el gobierno por la economía. Encierre en un círculo un número en cada una para señalar si está en favor o en contra.

- |                             |
|-----------------------------|
| 1. Muy en favor             |
| 2. En favor                 |
| 3. Ni en favor ni en contra |
| 4. En contra                |
| 5. Muy en desacuerdo        |

ENCIERRE EN UN CÍRCULO UN NÚMERO

a. Control oficial de los ingresos . . . . .	1	2	3	4	5	28/
b. Control oficial de los precios . . . . .	1	2	3	4	5	29/
c. Recortes de los gastos del gobierno . . . . .	1	2	3	4	5	30/
d. Financiamiento gubernamental de proyectos que creen empleos nuevos . . . . .	1	2	3	4	5	31/
e. Menos regulaciones oficiales de las empresas . . . . .	1	2	3	4	5	32/
f. Respaldo a las industrias para que desarrollen nuevos productos y tecnología . . . . .	1	2	3	4	5	33/
g. Respaldo a las industrias en problemas para proteger los empleos . . . . .	1	2	3	4	5	34/
h. Reducir la semana laboral para crear más empleos . . . . .	1	2	3	4	5	35/

11. La siguiente es una lista de varios rubros de gasto gubernamental. Indique si quisiera ver más gastos del gobierno en cada rubro. Recuerde que si dice "mucho más", podría ser necesario aumentar los impuestos para cubrir el gasto.

- |                              |
|------------------------------|
| 1. Gastar mucho más          |
| 2. Gastar más                |
| 3. Gastar lo mismo que ahora |
| 4. Gastar menos              |
| 5. Gastar mucho menos        |
| 6. Sin elección              |

ENCIERRE EN UN CÍRCULO UN NÚMERO

a. El ambiente . . . . .	1	2	3	4	5	8	36/
b. Salud . . . . .	1	2	3	4	5	8	37/
c. Policía y aplicación de la ley . . . . .	1	2	3	4	5	8	38/
d. Educación . . . . .	1	2	3	4	5	8	39/
e. Ejército y defensa . . . . .	1	2	3	4	5	8	40/
f. Beneficios de jubilación . . . . .	1	2	3	4	5	8	41/
g. Beneficios de desempleo . . . . .	1	2	3	4	5	8	42/
h. Cultura y artes . . . . .	1	2	3	4	5	8	43/

12. Si el gobierno tuviera que escoger entre mantener baja la inflación o reducir el desempleo, ¿a qué cree que deba concederle la mayor prioridad?

Mantener baja la inflación . . . . .	1	44/
Mantener bajo el desempleo . . . . .	2	
Sin elección . . . . .	8	

13. ¿Cree que los sindicatos laborales del país tienen mucho poder o muy poco?

Demasiado poder . . . . .	1	45/
Mucho poder . . . . .	2	
El poder correcto . . . . .	3	
Poco poder . . . . .	4	
Muy poco poder . . . . .	5	
Sin elección . . . . .	8	

**Figura 6.8**

Cuestionario de muestra (continuación)

14. En cuanto a las empresas e industrias, ¿tienen mucho poder o muy poco?						
Demasiado poder . . . . .	.1				46/	
Mucho poder . . . . .	.2					
El poder correcto . . . . .	.3					
Poco poder . . . . .	.4					
Muy poco poder . . . . .	.5					
Sin elección . . . . .	.8					
15. Y el gobierno federal, ¿tiene mucho poder o muy poco?						
Demasiado poder . . . . .	.1				47/	
Mucho poder . . . . .	.2					
El poder correcto . . . . .	.3					
Poco poder . . . . .	.4					
Muy poco poder . . . . .	.5					
Sin elección . . . . .	.8					
16. En general, ¿cómo diría que son los sindicatos para el conjunto del país?						
Excelentes . . . . .	.1				48/	
Muy buenos . . . . .	.2					
Buenos . . . . .	.3					
No muy buenos . . . . .	.4					
Nada buenos . . . . .	.5					
Sin elección . . . . .	.8					
17. ¿Cuál cree que deba ser la función del gobierno en cada una de estas industrias?						
1. Poseerlas 2. Controlar precios y ganancias pero no poseerlas 3. Ni poseerlas ni controlar precios ni ganancias 8. Sin elección						
<u>ENCIERRE EN UN CÍRCULO UN NÚMERO</u>						
a. Energía eléctrica . . . . .	1	2	3	8	49/	
b. Industria acerera . . . . .	1	2	3	8	50/	
c. Banca y seguros . . . . .	1	2	3	8	51/	
18. En conjunto, señale si debe o no ser la responsabilidad del gobierno...						
1. Definitivamente debe ser 2. Probablemente debe ser 3. Probablemente no debe ser 4. Definitivamente no debe ser 8. Sin elección						
<u>ENCIERRE EN UN CÍRCULO UN NÚMERO</u>						
a. Dar empleo a todo el que lo quiera . . .	1	2	3	4	8	52/
b. Controlar los precios . . . . .	1	2	3	4	8	53/
c. Dar atención médica a los enfermos . . .	1	2	3	4	8	54/
d. Proporcionar un nivel de vida decoroso a los ancianos . . . . .	1	2	3	4	8	55/

Figura 6.8

Cuestionario de muestra (continuación)

e. Dar a las industrias la ayuda que necesitan para crecer . . . . .	1	2	3	4	8	56/	
f. Proporcionar un nivel de vida decoroso a los desempleados . . . . .	1	2	3	4	8	57/	
g. Reducir las diferencias de ingresos entre ricos y pobres . . . . .	1	2	3	4	8	58/	
h. Dar ayuda económica a los estudiantes universitarios de familias de escasos recursos . . . . .	1	2	3	4	8	59/	
i. Proporcionar alojamiento decoroso a quienes no pueden costearlo . . . . .	1	2	3	4	8	60/	
19. ¿Qué tan interesado diría que está usted en la política?							
Muy interesado . . . . .	.1						61/
Apenas interesado . . . . .	.2						
Un poco interesado . . . . .	.3						
No muy interesado . . . . .	.4						
No interesado en absoluto . . . . .	.5						
Sin elección . . . . .	.8						
19. Los siguientes son otros rubros de gasto gubernamental. Indique si quisiera que el gobierno gastara más en cada rubro. Recuerde que si dice "mucho más", podría ser necesario aumentar los impuestos para cubrir el gasto.							
1. Gastar mucho más 2. Gastar más 3. Gastar lo mismo que ahora 4. Gastar menos 5. Gastar mucho menos 8. Sin elección							
<u>ENCIERRE EN UN CÍRCULO UN NÚMERO</u>							
a. Cuidado prenatal de las gestantes que no pueden pagarlo . . . . .	1	2	3	4	5	8 62/	
b. Atención médica para los niños de familias que no tienen seguro . . . . .	1	2	3	4	5	8 63/	
c. Programas preescolares como Head Start para los niños pobres . . . . .	1	2	3	4	5	8 64/	
d. Guarderías para los niños pobres . . . . .	1	2	3	4	5	8 65/	
e. Guarderías para todos los hijos de padres que trabajan . . . . .	1	2	3	4	5	8 66/	
f. Alojamiento para las familias pobres con hijos . . . . .	1	2	3	4	5	8 67/	
g. Servicios para los niños incapacitados o con enfermedades crónicas . . . . .	1	2	3	4	5	8 68/	
h. Prevención y tratamiento de la drogadicción de niños y jóvenes . . . . .	1	2	3	4	5	8 69/	
i. Programas de nutrición para niños y familias pobres, como cupones de alimentos y almuerzos escolares . . . . .	1	2	3	4	5	8 70/	
j. Servicios anticonceptivos para adolescentes . . . . .	1	2	3	4	5	8 71/	
MUCHAS GRACIAS POR CONTESTAR EL CUESTIONARIO							



---

## Puntos principales

- La operacionalización es una extensión del proceso de conceptualización.
- En la operacionalización, especificamos procedimientos empíricos concretos que darán por resultado mediciones de variables.
- La operacionalización es la última especificación de la manera en que reconoceremos los atributos de determinada variable en la realidad.
- Aunque la operacionalización como la estudiamos en el capítulo es esencial para la elaboración de cuestionarios, su lógica también es provechosa para otros modos de observación.
- Al determinar el margen de variación de una variable, no olvide considerar el concepto opuesto. ¿Será suficiente medir la religiosidad de "mucho" a "ninguna", o habría que proseguir para medir también la "antirreligiosidad"?
- Las mediciones nominales describen variables con atributos que simplemente son diferentes entre sí, como género. Las mediciones ordinales se refieren a variables con atributos que podemos ordenar según una progresión de mayor a menor. Un ejemplo es la variable prejuicios, compuesta de los atributos muy prejuicioso, un poco prejuicioso, ligeramente prejuicioso y nada prejuicioso en absoluto.
- Las medidas intervalares se refieren a las variables cuyos atributos no sólo están ordenados, sino también separados por una distancia uniforme. Un ejemplo es el CI.
- Las medidas de razón son las mismas que las intervalares, salvo que además se basan en un punto cero verdadero. La edad es un ejemplo de medida de razón, pues la variable contiene el atributo cero años.
- A veces, una variable puede medirse en varios niveles. Así, la edad que es potencialmente una medida de razón puede tratarse como intervalar,

ordinal y aun nominal. El nivel de medición más apropiado depende del propósito de la medición.

- Los cuestionarios ofrecen un método para recopilar datos al (1) formular preguntas a la gente, y (2) pedirles que concuerden o no con enunciados que manifiestan varios puntos de vista.
- Las preguntas pueden ser abiertas (los entrevistados dan sus propias respuestas) o cerradas (eligen de una lista fija de respuestas).
- Por lo regular, los reactivos cortos son mejores en los cuestionarios que los largos.
- En los cuestionarios hay que evitar reactivos y términos negativos, porque pueden confundir a los entrevistados.
- Los reactivos tendenciosos de los cuestionarios alientan a los entrevistados a responder de cierta manera o a respaldar determinados puntos de vista. Evítelos.
- La redacción de las preguntas y el formato de los cuestionarios influyen en la calidad de los datos reunidos.
- Las preguntas condicionadas son aquellas que contestan sólo algunos entrevistados según hayan respondido a preguntas anteriores.
- La matriz de preguntas es un formato eficaz para presentar varios reactivos que comparten la misma categoría de respuestas.
- El orden de los reactivos en el cuestionario puede influir en las respuestas dadas.
- Las instrucciones claras son importantes para obtener las respuestas adecuadas al cuestionario.
- La operacionalización comienza en el diseño del estudio y continúa durante todo el proyecto de investigación, incluyendo el análisis de datos.

---

## Preguntas y ejercicios de repaso

1. ¿Qué nivel de medición nominal, ordinal, intervalar o de razón describe a cada una de las variables siguientes?:

- a. Grupo étnico (blanco, negro, asiático, etc.)
  - b. Orden de llegada en una carrera (primer lugar, segundo, tercero, etc.)
  - c. Número de hijos en las familias
  - d. Habitantes de las naciones
  - e. Opiniones sobre la energía nuclear (aprobación firme, aprobación, desaprobación, desaprobación firme)
  - f. Región de nacimiento (noreste, sureste, etc.)
  - g. Orientación política (muy liberal, un poco liberal, un poco conservador, muy conservador)
2. Con cada una de las siguientes preguntas abiertas elabore una pregunta cerrada que pueda servir en un cuestionario.
    - a. ¿Cuál fue el ingreso total de su familia el año pasado?
    - b. ¿Qué opina del programa de transbordadores espaciales?
    - c. ¿Cuál es la importancia de la religión en su vida?
    - d. ¿Cuál fue su principal razón para asistir a la universidad?
    - d. ¿Cuál cree que es el principal problema que enfrenta su comunidad?
  2. Encuentre un cuestionario en internet (Sugerencia: Busque "cuestionario"). Haga una crítica positiva o negativa de por lo menos cinco de sus preguntas. No olvide citar la dirección (URL) del cuestionario y la redacción fiel de las preguntas que analizó.
  2. Visite la página en la red de la Encuesta Social Genera (<http://www.icpsr.ummich.edu/gss/subject/s-index.htm>). Identifique por lo menos tres preguntas en el libro de códigos que pudieran ser útiles para algún propósito de aplicación. Explique cómo aprovecharía las preguntas una empresa, gobierno o cliente no comercial.

---

## Proyecto de continuidad

Escriba 10 preguntas de un cuestionario que aborden las opiniones sobre la igualdad sexual. Presente las preguntas como aparecerían en el cuestionario, según los formatos ejemplificados en este capítulo.

---

## Lecturas adicionales

- Feick, Lawrence F., "Latent Class Analysis of Swvey Questions That Include Don't Know Responses", en *Public Opinion Quarterly* 53, núm. 4 (invierno de 1989): 525-47. Como lo indica este análisis, no se pueden significar muchas cosas.
- Fowler, Floyd J., Jr., *Improving Survey Questions: Design and Evaluation*, Thousand Oaks, Cal., Sage, 1995. Estudia la lógica de la consecución de información mediante preguntas y da numerosos lineamientos para ser eficaz. Ofrece varios ejemplos de preguntas que usted puede emplear.
- Miller, Delbert, *Handbook of Research Design in Social Measurement*, Newbury Park, Cal., Sage, 1991. Libro de referencia convincente. Este libro, en particular la parte 6, cita y describe una amplia variedad de mediciones operacionales de las primeras investigaciones sociales. En varios casos, se presentan los formatos de los cuestionarios. Aunque la calidad de las ilustraciones es dispareja, son ejemplos excelentes de las variaciones posibles.
- Schwartz, Norman *et al.*, "Rating Scales: Numeric Values May Change the Meaning of Scale Labels", en *Public Opinion Quarterly* (invierno de 1991): 570-82. Unos investigadores alemanes experimentaron con varios formatos de escalas y descubrieron que los cambios influyen en las respuestas.
- Sheatsley, Paul F., "Questionnaire Construction and Item Writing", en Peter H. Rossi, James D. Wright y Andy B. Anderson (comps.), *Handbook of Survey Research*, Nueva York, Academic Press, 1983, pp. 195-230. Excelente examen del tema por un experto en el campo.
- Smith, Eric R. A. N., YPeverill Squire, "The Effects of Prestige Names in Question Wording", en *Public Opinion Quarterly* Q4 (verano de 1990): 97-116. Los nombres de prestigio no sólo influyen en las respuestas generales dadas a un cuestionario de encuesta, sino que también tienen un efecto en cosas como la correlación entre educación y el número de respuestas "no sé".
- Tourangeau, Roger *et al.*, "Carryover Effects in Attitudes Surveys", en *Public Opinion Quarterly* 53 (invierno de 1989): 495-524. En una encuesta telefónica a 1100 entrevistados, los autores formularon seis preguntas objetivos variando las preguntas inmediatamente anteriores. Encontraron diferencias sustanciales.