

- **Babbie**, E. (1996) *Manual para la práctica de la investigación social*, (Bilbao: Desclée De Brouwer) Capítulo 5: "Conceptualización y medición".

5

Conceptualización y medición

INTRODUCCIÓN

Este es el primero de los tres capítulos en los que se aborda el proceso de cambio desde las ideas vagas sobre lo que quieres estudiar, hasta ser az de reconocerlo y medirlo en el mundo real. En este capítulo se hablará de la idea general de *conceptualización*, que sienta la base para el comentario sobre la *operacionalización* en el Capítulo 6. Concluiremos en el Capítulo 7, los temas de los que hablamos en los Capítulos 5 y 6, conociendo tipos de medidas más complejas.

Quisiera empezar el capítulo con un ataque frontal a la preocupación oculta que algunas veces tenemos sobre si es posible medir las cosas cotidianas: amor, odio, prejuicios, radicalismo, alienación y cosas como esas. La respuesta es sí, pero me va a llevar algunas páginas explicarlo. Una vez que hayamos visto que se puede medir todo lo que existe, nos centraremos en los pasos necesarios para hacerlo.

MEDIR TODO LO QUE EXISTE

Creo que es posible que pueda tener algunas reservas sobre la habilidad de la ciencia para medir los aspectos realmente importantes de la existencia social humana. Puede que haya leído artículos de investigación que tratan de temas como el liberalismo, la religión o los prejuicios, y puede que no esté satisfecho con la manera en que los investigadores han medido lo que estaban estudiando. Puede que haya pensado que eran demasiado superficiales, que olvidaban los aspectos más importantes. Puede que midieran la *religiosidad* como el número de veces que una persona acude a la iglesia, o puede que midieran el *liberalismo* por lo que se votaba sólo en unas elecciones. Seguramente su insatisfacción haya aumentado si encuentra que a usted mismo se le ha clasificado mal con ese sistema de medida. A menudo tenemos esa experiencia.

O, puede que haya buscado en el diccionario la definición de una palabra como *compasión* y después de encontrarla se diga a sí mismo: "significa mucho más que eso". De hecho, siempre que busque la definición de algo que ya entiende bien, probablemente pueda ver maneras en que otros podrían equivocarse el término si sólo cuentan con esa definición.

Anteriormente, decía que uno de los dos pilares de la ciencia es la *observación*. Como esta palabra puede tener un matiz de actividad casual y pasiva, los científicos suelen utilizar en su lugar la *medición*, que significa observaciones deliberadas y cuidadosas del mundo real con el propósito de describir objetivos y acontecimientos en términos de atributos que componen una variable. Si la variable estudiada es la *afiliación a un partido político*, podríamos consultar el censo electoral para saber si las personas que estamos estudiando se inscriben como demócratas o republicanos. De esta manera, habríamos medido la afiliación a un partido político.

¿Cómo Conocemos?

Para demostrarle que los científicos sociales pueden medir todo lo que existe, me gustaría que imaginase que estamos discutiendo sobre un tema. Yo escribiré el guión, pero usted puede sustituir cualquier frase por otra siempre que lo desee.

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

USTED: ¡Bah! ¡No lo creo!

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

USTED: Bien, a ver cómo mide el prejuicio.

YO: Buena elección. Ahora, no vamos a perder el tiempo intentando medir algo que no existe. Así que, dígame si existe.

USTED: Sí, claro que existe. Todo el mundo sabe que el prejuicio existe, todos lo conocemos. YO: También hubo un tiempo en que todo el mundo pensó que la tierra era plana, pero explíqueme cómo sabe que realmente existe el prejuicio.

USTED: Lo he visto actuar.

YO: ¿Qué ha visto que pruebe que existe el prejuicio?

USTED: En una ocasión un hombre de negocios me dijo que nunca contrataría a una mujer para un puesto ejecutivo porque pensaba que todas las mujeres eran volubles e irracionales. ¿Qué le parece?

YO: Muy bien. Eso me suena a prejuicio, así que vamos a asumir que existe el prejuicio. Ahora voy a medirlo. ¿Preparado?

USTED: Preparado.

YO: Usted y yo vamos a pasear tranquilamente por la zona de negocios, hablando a los hombres de negocios sobre los contratos. Siempre que uno de ellos nos diga que nunca contrataría a una mujer para un puesto ejecutivo porque piensa que todas las mujeres son volubles e irracionales, lo contaremos como un caso de prejuicio. Siempre que no nos digan eso, consideraremos la conversación como un caso de no prejuicio. Cuando terminemos, podremos clasificar a todos los hombres de negocios con los que hayamos hablado como personas con prejuicios o sin prejuicios.

USTED: ¡Un momento! Esa no es una manera buena de medir los prejuicios. Así dejamos de lado un montón de prejuicios. Sólo mediremos al prejuicio evidente a la hora de contratar mujeres.

YO. Comprendo lo que quiere decir. Pero su comentario también significa que la situación que describió anteriormente sólo prueba que existe un prejuicio evidente para contratar mujeres. Haríamos mejor en reconsiderar si existe el prejuicio, ¿no?

USTED: Claro que existe. Yo sólo estaba dándole un ejemplo de prejuicio, pero hay cientos de ejemplos.

YO: Déme uno que lo pruebe.

USTED: Muy bien, vamos a intentarlo con otro. Estaba en el campus la otra noche, y dos individuos, uno blanco y otro afroamericano, estaban discutiendo sobre política. Finalmente, el blanco estaba tan enfadado que empezó a utilizar un desagradable lenguaje racista y decía: "Tendrían que devolver a África a toda la gente como tú". ¿Le parece suficiente este prejuicio?

YO: Me gusta. Parece probar que existen los prejuicios, así que estoy preparado otra vez para medir el prejuicio. Nos dividiremos y empezaremos a recorrer el campus todas las noches. Agudizaremos el oído y escucharemos a personas blancas que usen desagradables epítetos raciales diciendo: "Tendrían que devolver...".

USTED: ¡Espere! Ya sé por dónde va y no puede ser. Una persona que dice eso tendrá prejuicios, pero vamos a clasificar a un montón de personas con prejuicios y sin ellos, sólo porque no han usado epítetos raciales.

YO: Con lo cual volvemos a mi pregunta original. ¿Existe realmente el prejuicio o me está tomando el pelo?

USTED: ¡Sí que existe!

YO: Bueno, ya no sé si me lo creo. Primero me quiere convencer de que existen hombres de negocios que discriminan a las mujeres en los contratos, porque lo vió y yo le he creído. Después, me convence de que hay personas que insultan a los afroamericanos y les dicen que deberían volver a África. Pero ya no estoy tan seguro de que exista el prejuicio. Me gustaría estar seguro de lo que es, para poder demostrarle que puedo medirlo. Aunque, para ser honesto, estoy empezando a dudar de que realmente exista. Lo que quiero decir es: ¿alguna vez ha visto un prejuicio?, ¿de qué color es?, ¿cuánto pesa?, ¿dónde está?

USTED: ¿De qué está hablando?

Llegados a este punto del diálogo como ya habrá adivinado, se demuestra que el *prejuicio no existe*. No sabemos a qué se parece un prejuicio, qué tamaño o color tiene. Nadie ha tocado nunca un prejuicio, ni se ha subido en él, pero hablamos mucho sobre ello. Así es como son las cosas.

A lo largo de nuestra vida observamos muchas cosas y sabemos que son reales gracias a nuestras observaciones. Oímos hablar de otras muchas cosas que otras personas dijeron que habían observado y pensamos que existen también. Por ejemplo, alguien contó que había visto un linchamiento y lo describió con gran detalle.

Si poseemos algo más de experiencia, podemos darnos cuenta de que las personas que participan en los linchamientos es probable que también insulten a los afroamericanos. Muchos de ellos querían que las "mujeres se queden en su sitio". Finalmente, comenzaremos a sentir que en el mundo hay cierto tipo de personas que tienen esas tendencias. Cuando discutimos sobre la gente que conocemos, algunas veces definimos a alguien en términos de tales tendencias. Solemos decir que una persona es "de esos que participan en linchamientos, insultan a los afroamericanos y no quieren que una mujer ocupe un puesto ejecutivo". Sin embargo, un rato después, parece un poco exagerado decir eso y se le ocurre una brillante idea: "Vamos a usar la palabra *prejuicio* para definir a personas de estas características. Podremos usar ese término aunque no manifiesten todos esos comportamientos, bastará con que haya muchos de ellos."

Si queremos ser eficaces, estoy básicamente de acuerdo y acepto continuar con ese sistema. De ahí procede el *prejuicio*. Nunca ha existido en realidad, nunca lo hemos visto, sólo hemos inventado una definición corta para hablar a espaldas de la gente. Realmente el *prejuicio* es meramente un término que acordamos usar en la comunicación: un nombre que utilizamos para representar una colección completa de fenómenos aparentemente relacionados que hemos observado en el curso de la vida. Cada uno de

nosotros ha desarrollado su propia imagen mental sobre lo que representa en general y tiene en común el conjunto de fenómenos que hemos observado.

Cuando nombra la palabra *prejuicio*, sé que evoca una imagen mental en su cerebro, de la misma manera que me ocurre a mí. Es como si tuviéramos muchos archivos en nuestras cabezas que contienen miles de hojas de papel, y cada hoja tiene una etiqueta en la parte superior derecha. Una de las hojas de su archivo contiene el término *prejuicio* y yo también la tengo. En su hoja se encuentran todas las cosas que le han dicho sobre el *prejuicio* y todos los ejemplos que ha observado. En una hoja está todo lo que me han contado sobre el *prejuicio*, más todo lo que he observado.

Concepciones y Conceptos

El término técnico para esas imágenes mentales, hojas de papel dentro de nuestros archivos mentales, es *concepción*. Cada una de las hojas es una *concepción*. Pero esas imágenes no pueden comunicarse directamente. No hay manera de revelar directamente lo que está escrito en mi mente. Por lo tanto, utilizamos los *términos* escritos en la esquina superior derecha para poder comunicar nuestras concepciones y lo que observamos en relación a esas concepciones.

Supongamos que voy a conocer a alguien que se llama Pat a quien usted ya ha tratado. Le preguntaré cómo es Pat. Imaginemos que usted ha visto a Pat ayudando a niños perdidos a encontrar a sus padres y devolviendo un pajarito a su nido. Usted mismo acompaña a Pat a llevar pavos el día de Acción de Gracias a familias pobres y visita a niños que están en el hospital el día de Navidad. Ha visto llorar a Pat al ver una película de una madre que abandona a su hijo. Cuando busca en su fichero mental, encuentra todos o la mayoría de estos fenómenos grabados en la hoja cuya etiqueta de la parte superior derecha dice: *compasiva*. Repasa el resto de lo que está escrito en la hoja y averigua que describe con bastante exactitud a Pat. Así que nos dice: "Pat es *compasiva*".

Ahora, yo busco en mi fichero mental la hoja marcada con la palabra *compasiva*. Reviso lo que está escrito en ella y digo: "Ah, está bien". Ya sé cómo es Pat, pero mis expectativas al respecto son un reflejo de mi ficha, no de la suya. Más tarde, cuando conozco a Pat puedo darme cuenta de que mis propias experiencias personales corresponden a lo registrado en la ficha *compasiva* y compruebo que usted tenía razón. O mis observaciones de Pat puede que entren en contradicción con mi ficha, y entonces le diré que no creo que Pat sea una persona muy *compasiva*. Si ocurriese esto último, empezaremos a contrastar lo que pensamos.

Usted me dice: "En una ocasión vi que Pat lloraba con una película sobre una madre que se enfrentaba a un montón de adversidades para salvar y proteger a su hijo". Entonces busco en mi ficha de *compasiva* y no encuentro nada que se asemeje a eso. Pero buscando un poco más en mi fichero, localizo este tipo de fenómeno en la hoja etiquetada como *sentimental*. Por tanto, replico: "Eso no es *compasión*, es *sentimentalismo*".

Para reforzar mi argumento, le digo que he visto a Pat rehusando dar dinero a una organización que se dedica a salvar a las ballenas en peligro de extinción. "Este hecho representa una falta de *compasión*", explico. Usted busca en sus fichas y encuentra que salvar a las ballenas está en una hoja marcada con las palabras *activismo medioambiental*, y me lo dice. Finalmente, comparamos los registros de nuestras

respectivas fichas con la palabra *compasiva* y descubrimos imágenes mentales bastante diferentes para representar ese término.

Hablando en sentido amplio, el lenguaje y la comunicación trabajan sólo en la dimensión que usted y yo tenemos coincidencias en los registros que corresponden a nuestras fichas mentales. Las similitudes que encontramos en esas fichas representan el acuerdo que existe en la sociedad en que vivimos. Cuando crecimos, nos hablaron de aproximadamente las mismas cosas al explicarnos un término en particular. Los diccionarios son una manera de formalizar los acuerdos de nuestra sociedad sobre tales términos. Cada uno de nosotros, por tanto, se formó su imagen mental correspondiente a esos acuerdos, pero como todos tenemos experiencias y observaciones diferentes, no hay dos personas que terminen tendiendo exactamente el mismo conjunto de registros en cualquiera de las hojas de su fichero. Volviendo a la afirmación que habíamos hecho al principio de este capítulo, podemos medir cualquier cosa que sea real. Por ejemplo, podemos medir si Pat realmente devuelve el pájaro a su nido, visita el hospital en Navidad, llora con la película o no quiere contribuir a la salvación de las ballenas. Todos estos hechos existen, luego podemos medirlos. Pero, ¿es Pat en realidad una persona *compasiva*? No se puede responder a esa pregunta: no podemos medir la compasión en ese sentido, porque la compasión no existe de la misma forma en que existen los acontecimientos que acabo de describir.

La *compasión* como *término* sí existe. Podemos contar el número de letras que tiene y acordar que son nueve, así como que consta de tres sílabas y comienza por la letra C. En definitiva, se pueden medir aspectos de la palabra que son reales.

Algunos aspectos de nuestras concepciones son también reales, por ejemplo, que usted tiene una imagen mental asociada con el término *compasión*. Cuando un profesor de primaria pregunta a la clase cuántos saben lo que significa *compasión*, se puede contar a todos los que han levantado la mano.

La presencia de determinados registros en las fichas que tienen una determinada etiqueta es también real y se puede medir. Podríamos medir cuántas personas dan o no dinero para salvar a las ballenas, según su concepto de compasión. Lo único que no podemos medir es qué significa realmente la palabra compasión, porque la compasión no es real. La compasión existe sólo en forma de acuerdos que tenemos sobre cómo usar el término a la hora de comunicar cosas que son reales.

En este contexto, Abraham Kaplan (1964) distingue tres tipos de cosas que miden los científicos. Las primeras son *directamente observables*: todo lo que podemos observar simple y directamente, como el color de una manzana o la respuesta elegida en un cuestionario. Lo que es *indirectamente observable* requiere "observaciones más sutiles, complejas o indirectas" (1964: 55). Vemos que una persona ha marcado la casilla *femenino* en un cuestionario y así tenemos una observación indirecta de su sexo. Los libros de historia o las conferencias nos proporcionan observaciones indirectas de acciones sociales pasadas. Finalmente, los *constructos* son creaciones teóricas basadas en observaciones, pero que no pueden observarse directa o indirectamente. El cociente intelectual es un buen ejemplo. Se construyó matemáticamente a partir de observaciones de respuestas dadas a un gran número de preguntas de una prueba de C. I. Otras medidas compuestas de las que trataremos en el Capítulo 7 son más ejemplos de constructos.

Kaplan (1964: 49) define el *concepto* como "una familia de concepciones". Un concepto es, según Kaplan, un constructo. El concepto *compasión*, por tanto, es un constructo creado a partir de su concepción sobre él, de la mía y de las concepciones de todos aquellos que alguna vez lo han utilizado. No puede observarse directa o indirectamente porque no existe. Nosotros lo hemos creado.

Conceptualización

La comunicación diaria se produce a través de un sistema vago y general de acuerdos sobre el uso de los términos. Habitualmente, las personas no comprenden exactamente lo que queremos comunicar, pero captan su sentido general. Aunque usted y yo no estemos de acuerdo completamente en el uso del término *compasivo*, probablemente esté seguro de que Pat no va a extender sus alas para volar. El precio que tenemos que pagar por nuestras imprecisiones es un amplio rango de malos entendidos y conflictos - desde interpersonales, hasta internacionales, aunque de algún modo conseguimos solventarlos. Sin embargo, la ciencia aspira a más y no puede operar en un contexto tan impreciso.

La *conceptualización* es el proceso a través del cual especificamos con precisión lo que queremos decir cuando utilizamos un término en particular. Supongamos que queremos averiguar, por ejemplo, si las mujeres son más compasivas que los hombres. Me imagino que la mayoría pensamos que sí, pero sería más interesante averiguar si es verdad. No podemos estudiar significativamente la pregunta, ni ponernos de acuerdo en la respuesta sin trabajar en la precisión del significado del término. Se trabaja sobre acuerdos en el sentido que nos permiten estudiar la pregunta. No necesitamos estar de acuerdo o incluso pretender estar de acuerdo en que podría emplearse una especificación en particular.

Indicadores y Dimensiones

El producto final de este proceso de conceptualización es la especificación de un conjunto de *indicadores* que tenemos en mente, y que indican la presencia o ausencia del concepto que estamos estudiando. Por tanto, podemos estar de acuerdo en utilizar las visitas en Navidad a los niños que están en el hospital como un indicador de la compasión. Devolver los pájaros a sus nidos puede ser otro indicador, y así sucesivamente. Si las unidades de análisis de nuestro estudio son los individuos, entonces podríamos observar la presencia o ausencia de cada indicador en cada persona que forme parte del estudio. Incluso podríamos llegar más lejos y añadir el número de indicadores de compasión observados en cada individuo. Podríamos acordar diez indicadores específicos, por ejemplo, y encontrar seis de ellos cuando estudiamos a Pat, tres en John y nueve en Mary, etc.

Volviendo a nuestra pregunta original, podríamos calcular que las mujeres del estudio tendrían una media de 6,5 indicadores de compasión y los hombres una media de 3,2. Por lo tanto, podríamos llegar a la conclusión, en base a las diferencias entre los dos grupos, de que las mujeres son, en conjunto, más compasivas que los hombres. Normalmente, las cosas no son tan sencillas.

A menudo, cuando tomamos en serio los conceptos e intentamos especificar lo que significan, descubrimos los desacuerdos e inconsistencias. No sólo usted y yo disentimos, sino que es muy probable que cada uno de nosotros encuentre una buena dosis de confusión en sus propias imágenes mentales. Si se para un momento a pensar lo que *usted* entiende por compasión, probablemente descubra que su imagen contiene varias *clases* de compasión. Los registros de su fichero pueden combinarse en grupos y subgrupos, e incluso puede encontrar varias estrategias diferentes para realizar las combinaciones. Por ejemplo, se podrían agrupar los registros en sentimientos y acciones.

El nombre técnico que se les da a estas agrupaciones es *dimensión*: un aspecto o faceta específicos de un concepto. Por tanto, podríamos hablar de la "dimensión sentimental" de la compasión y de la "dimensión activa" de la compasión. Con un esquema diferente de agrupación, podríamos distinguir entre "compasión por los seres humanos" y la "compasión por los animales". O, la compasión podría centrarse en ayudar a las personas a ser y tener lo que quieren. Todavía podríamos distinguir entre la "compasión como misericordia" de la "compasión como pena".

Sería posible subdividir el concepto de compasión en función de varios grupos de dimensiones diferentes. Especificar dimensiones e identificar los indicadores para cada una de estas dimensiones son las dos partes de la conceptualización.

Especificar las distintas dimensiones de un concepto, suele facilitar el camino para comprender mejor lo que estamos estudiando. Podríamos observar, por ejemplo, que las mujeres son más compasivas en términos de sentimientos, y los hombres más compasivos en términos de acciones, o viceversa. Observe que si éste fuera el caso, no seríamos capaces de decir si los hombres o las mujeres son realmente más compasivos. De hecho, nuestra investigación habría demostrado que no hay una respuesta sencilla a la pregunta.

Los Indicadores Intercambiables

Supongamos, por un momento, que usted y yo hemos elaborado una lista de 100 indicadores para el concepto *compasión* y sus diferentes dimensiones. Supongamos también que no estamos de acuerdo en qué indicadores proporcionan la evidencia más clara de la compasión o su ausencia. En el caso de que sí estemos de acuerdo en algunos, centraríamos nuestra atención en ellos y probablemente acordaríamos la pregunta a la que responden. Pero si no estamos de acuerdo en ninguno de los posibles indicadores, todavía es posible llegar a un consenso respecto a si son los hombres o las mujeres más compasivos.

Si discrepamos completamente en el valor de los indicadores, una solución sería estudiarlos todos. Ahora, supongamos que las mujeres resultan ser más compasivas que los hombres en los 100 indicadores, es decir, en todos los que usted había elegido y en todos los míos. Entonces, podríamos estar de acuerdo en que las mujeres son más compasivas que los hombres, incluso aunque todavía discrepemos en qué significa la compasión en general.

Los *indicadores intercambiables* implican que si varios indicadores diferentes representan todos ellos, en cierto grado, al mismo concepto, entonces todos se comportarán de la misma manera que se comportaría el concepto si fuese real y pudiese ser observado. Por tanto, si las mujeres son en general más compasivas que los hombres,

deberíamos ser capaces de observar esa diferencia utilizando cualquier medida razonable de la compasión.

Ahora ya conoce la lógica fundamental de la conceptualización y la medición. La discusión que viene a continuación en este capítulo y el siguiente es principalmente una profundización en lo que acabamos de tratar. Sin embargo, antes de adentrarnos en elaboraciones más técnicas del marco fundamental, sería útil aclarar más tópicos generales.

En primer lugar, sé que las discusiones anteriores pueden no encajar exactamente con su conocimiento previo del significado de términos como *prejuicio y compasión*. Tendemos a comportarnos en la vida diaria como si tales términos tuvieran significados reales y esenciales. Por eso, en la siguiente subsección, me gustaría comentar brevemente cómo llegaríamos a este entendimiento.

En segundo lugar, para que toda esta discusión no vaya a crear una anarquía en el significado de las palabras, describiré algunas de las maneras en las que los científicos han ordenado la confusión con el fin de proporcionar coherencia, acuerdo y unanimidad en el significado de los términos. Debiera extraer de esta última discusión un recuperado sentido del orden -pero basado en una comprensión consciente y no en una aceptación casual del uso común.

La Confusión sobre las Definiciones y la Realidad

Resumiendo brevemente, nuestros conceptos se derivan de imágenes mentales (concepciones) que son la suma de grupos de observaciones y experiencias parecidas. Aunque las observaciones y experiencias son reales, nuestros conceptos son sólo creaciones mentales. Los términos asociados con los conceptos son meras invenciones creadas para poder sentir y comunicarnos. La palabra *prejuicio* es un ejemplo. En el fondo esta palabra sólo es un conjunto de letras y no tiene ningún significado intrínseco.

Sin embargo, muy a menudo, caemos en la trampa de creer que esos términos tienen significados reales. Ese peligro aumenta cuando empezamos a tomar los términos en serio e intentamos utilizarlos con precisión. Y el peligro todavía aumenta más con la presencia de expertos que parecen saber más que usted sobre lo que significan realmente los términos. Es muy fácil ceder ante la autoridad de los expertos en tal situación.

Una vez que hemos asumido que los términos tienen significados reales, comenzamos con la ardua labor de descubrir cuáles son esos significados reales y en qué consiste una medida genuina de ellos. La Figura 5.1 ilustra ese proceso. Sumamos mentalmente las observaciones reales porque las sumas son convenientes. Sin embargo, resultan tan convenientes que empezamos a pensar que son reales. El proceso de considerar como cosas reales las que no lo son, se llama materialización y la materialización de conceptos en la vida diaria es muy común.

Crear un Orden Conceptual

El diseño y ejecución de la investigación social requiere aclarar la confusión entre conceptos y realidad. Para conseguirlo, los pensadores y los científicos han encontrado útil distinguir tres clases de definiciones: *real, normal y operacional*. La primera refleja la materialización de los términos, y como nos advierte Carl G. Hempel:

Una definición "real", según la lógica tradicional, no es una estipulación que determina el significado de alguna expresión, sino una representación de la "naturaleza esencial" o de los "atributos esenciales" de alguna entidad. Sin embargo, la noción de la naturaleza esencial es tan vaga que convierte en inútil esta caracterización para una investigación rigurosa.

(1952: 6)

La especificación de conceptos en la investigación científica depende de definiciones nominales y operacionales. Una definición *nominal* es la que se *asigna* a un término. En medio del desacuerdo y confusión que se producen respecto a lo que un término significa realmente, los científicos especifican una definición operativa para poder investigar. Si queremos examinar el estatus socioeconómico (ESE), por ejemplo, podemos simplemente especificar que vamos a abordarlo como una combinación de logros educativos e ingresos económicos. En esta decisión sobre la definición hemos dejado fuera muchos otros aspectos posibles del ESE: estatus profesional, dinero en el banco, propiedades, linaje, estilo de vida, etc.

La especificación de definiciones nominales centra nuestra estrategia observacional, pero no nos permite observar. El siguiente paso consiste en especificar exactamente lo que vamos a observar, cómo lo vamos a hacer y qué interpretaciones vamos a dar a las posibles observaciones. A todo ello se le denomina definición operacional de concepto -una definición que especifica con precisión cómo se va a medir el concepto. Hablando en sentido estricto, una definición operacional es una descripción de las "operaciones" que se realizarán para medir un concepto.

Continuando con el caso ESE, podríamos decidir hacer dos preguntas a las personas que estamos estudiando:

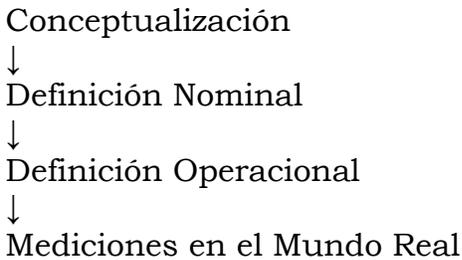
1. ¿Cuáles han sido los ingresos totales de la familia durante los últimos doce meses?

2. ¿Cuál fue su nivel educativo alcanzado?

En este momento, probablemente queramos especificar un sistema para categorizar las respuestas que nos vayan dando. Para el nivel de ingresos, podríamos usar categorías como "menos de 5.000\$" o "entre 5.000 y 10.000\$". También podríamos agrupar en categorías el nivel educativo. Finalmente especificaríamos la manera en que las respuestas de las personas a estas dos preguntas se combinarían para crear una medida de ESE. En el Capítulo 7, sobre la construcción de un índice y una escala, se presentarán algunos métodos para hacerlo.

Al final, habríamos creado una definición de ESE viable y con la que poder trabajar. Habrá quienes puedan no estar de acuerdo con nuestra conceptualización y operacionalización, pero la definición tendría una virtud científica esencial: sería absolutamente específica y sin ambigüedad. Incluso aunque alguien no esté de acuerdo con nuestra definición, podría interpretar los resultados de nuestra investigación, porque estaría claro lo que significaba para nosotros el término ESE -reflejado en el análisis y conclusiones-.

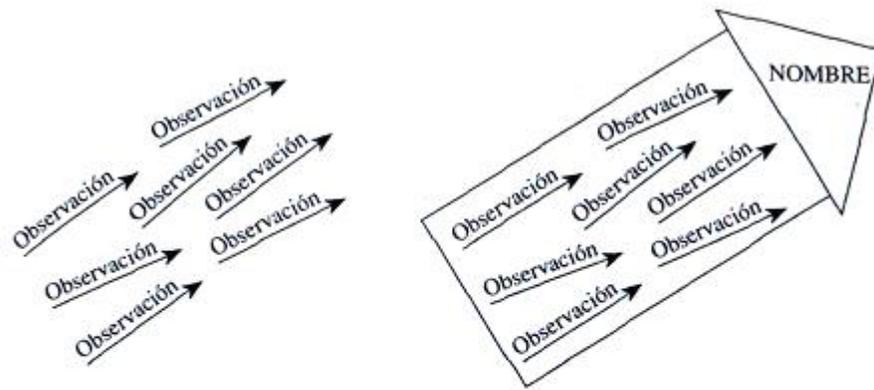
A continuación mostramos un diagrama de las etapas de medición desde el sentido vago que le damos a un término, hasta las mediciones específicas de un estudio científico:



Un Ejemplo de Conceptualización

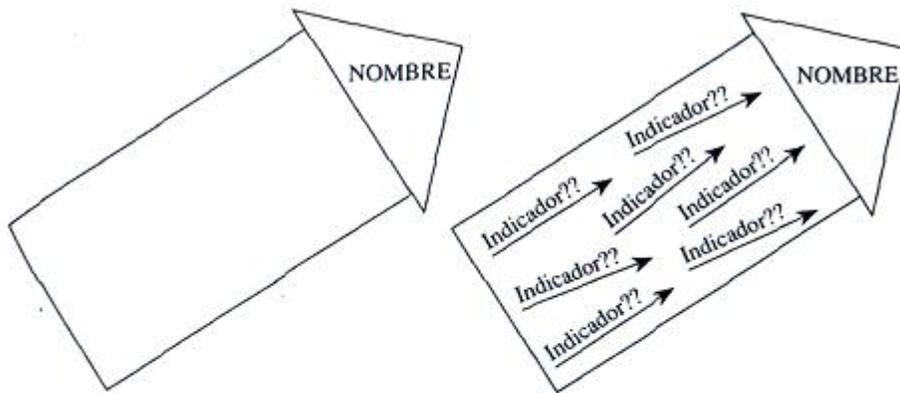
Quisiera retomar la anterior discusión junto con una breve historia de un concepto científico-social. Los investigadores que estudian los desórdenes urbanos suelen interesarse por el papel que desempeñan los sentimientos de "impotencia". Los científicos sociales a veces utilizan el término *anomía* en este contexto. Este término fue introducido por primera vez en las ciencias sociales por Emile Durkheim, el gran sociólogo francés, en *Suicidio*, su clásico estudio de 1897.

Durkheim escribió un trabajo analítico excepcional, utilizando solamente las publicaciones oficiales sobre los índices de suicidio en diferentes regiones y países. Para determinar los efectos de la religión en esta conducta, comparó los índices de suicidio de países predominantemente protestantes, con los predominantemente católicos, regiones protestantes de países católicos con regiones católicas de países protestantes, etc. Para determinar los posibles efectos del clima, comparó los índices de suicidio de países y regiones del norte y sur, así como en diferentes meses y estaciones del año. Así, pudo extraer conclusiones sobre el supremo acto individual y personal sin disponer de ningún dato sobre la implicación del individuo.



1. Muchas de nuestras observaciones en la vida diaria parecen tener algo en común. Sabemos que representan algo más general que el contenido simple de una observación sencilla. Es más, encontramos útil comunicar el concepto general.

2. No es conveniente seguir describiendo todas las observaciones específicas siempre que queramos comunicar el concepto general que parecen tener en común, así que damos un nombre al concepto general, queriendo abarcar todas las observaciones específicas que tienen en común.



3. Cuando comunicamos el concepto general, utilizando su término, empezamos a pensar que el concepto es alguna *cosa* que existe realmente, y no sólo una referencia que resume varias observaciones concretas.

4. La creencia de que el concepto en sí mismo es real resulta una ironía. Ahora comenzamos a discutir y debatir si las observaciones específicas son "realmente" indicadores suficientes del concepto.

Figura 5.1. El Proceso de la Trampa Conceptual

A un nivel más general. Durkheim sugirió que el suicidio también reflejaba el grado en que los acuerdos sociales eran claros y estables. Haciendo ver que las épocas de sublevaciones y cambios suelen provocar en los individuos graves dudas sobre lo que se espera de ellos. Durkheim sugirió que estas dudas causan confusión, ansiedad e incluso autodestrucción. Para descubrir esta condición social de falta de normas, Durkheim eligió el término *anomia*. Es importante saber que él no se inventó dicha palabra. Utilizada en alemán y francés, significaba literalmente *sin ley* y el término inglés "anomy"

se utilizó durante al menos tres siglos antes de Durkheim para expresar el desprecio por la ley divina. Sin embargo, Durkheim creó la anomía como concepto científico-social.

En los años que siguieron a la publicación de *Suicidio*, los científicos sociales encontraron que la anomía era un concepto útil y muchos ampliaron la utilización que le había dado Durkheim. Robert Merton, en un artículo clásico titulado "Estructura social y anomía" (1938), llegó a la conclusión de que este término era el resultado de la disparidad entre las metas y los significados prescritos por una sociedad. El éxito económico, por ejemplo, es una meta ampliamente compartida en nuestra sociedad, sin embargo no todos los individuos poseen los recursos para llegar a niveles aceptables. Según Merton, un énfasis en la meta en sí misma produce ausencia de normas porque quienes se niegan a llegar a la riqueza por el buen camino, buscan conseguirla a través de acciones ilegítimas. Por lo tanto, la discusión de Merton podría considerarse un avance en la conceptualización del concepto anomía.

Aunque Durkheim originalmente intentó que el concepto de anomía fuese una característica de las sociedades, como hizo Merton después de él, otros científicos sociales lo han utilizado para describir a los individuos. (Para aclarar esta distinción, algunos estudiantes han decidido emplear el término anomía en su sentido original, de tipo social y hacerlo en referencia a características individuales). Entonces, en una sociedad determinada, algunos individuos experimentan anomía y otros no. Elwin Powell, escribió 20 años después de Merton la siguiente conceptualización de la anomía (empleando el término "anomía") como una característica de los individuos:

Cuando el final de una acción se convierte en contradictorio, inaccesible o insignificante, surge la condición de anomía. Caracterizada por una pérdida general de orientación y acompañada por sentimientos de "vacío" y apatía, la anomía puede concebirse simplemente como una falta de sentido.

(1958: 132)

Powell continuó sugiriendo que había dos clases distintas de anomía y examinó cómo las dos surgían de diferentes experiencias ocupacionales para terminar, algunas veces, en suicidio. Sin embargo, en este estudio, Powell no midió la anomía per se; estudió la relación entre suicidio y ocupación, haciendo inferencias de los dos tipos de anomía. Por eso, el estudio no proporciona una definición operativa de la anomía, sólo una mayor conceptualización.

Muchos otros investigadores han ofrecido definiciones operativas, pero una de ellas sobresale por encima de todas las demás. Dos años antes de que apareciera el artículo de Powell, Leo Srole (1956) publicó un conjunto de cuestionarios que según él proporcionaban una buena medida de la anomía como experiencia individual. Consiste en cinco afirmaciones que los sujetos deben responder si están de acuerdo o en desacuerdo.

1. En contra de lo que algunos dicen, el hombre medio está empeorando.
2. Es bastante ingenuo traer niños al mundo con un futuro tan negro como el que tenemos.
3. Hoy en día una persona tiene que vivir mucho más el presente y dejar que el mañana siga su propio curso.
4. Actualmente una persona no sabe realmente con quién puede contar.

5. No sirve de nada escribir a las autoridades públicas porque en realidad no están interesadas en los problemas del hombre medio.

(1956: 713)

En las décadas que siguieron a su publicación, la escala Srole se convirtió en un asunto principal de las investigaciones de los científicos sociales. Es probable que encuentre que se ha utilizado esta particular operacionalización de la anomía en muchos de los proyectos de investigación publicados en revistas académicas (El profesor Srole trata de todo ello en el libro "Los orígenes de la Anomía", en donde discute la historia de su investigación en este área.

Si usted está interesado en el concepto de anomía, le presentamos algunos datos de un informe nacional dirigido por el Centro Nacional de Investigación de Opinión de la Universidad de Chicago en 1990, como parte de sus avances en Investigación Social General (Davis y Smith, 1990: 231-232). Para cada uno de los ítems mostrados, he indicado el porcentaje de quienes están de acuerdo entre aquellos que tienen una opinión.

	Porcentaje de quienes están de acuerdo
En contra de lo que algunos dicen, el hombre medio está empeorando.	58%
Es bastante ingenuo traer niños al mundo con un futuro tan negro como el que tenemos.	36%
La mayoría de las autoridades públicas (personas que ocupan un cargo público) no están realmente interesadas en los problemas del hombre medio.	69%

He presentado esta historia abreviada de la anomía como concepto científico-social por varias razones. Primera, ilustran el proceso a través del cual los conceptos, generales se convierten en medidas operativas, aunque no quisiera que pensase que el concepto de anomía está resuelto de una vez por todas. Seguramente, los estudiantes continuarán reconceptualizando y reoperacionalizándolos durante los próximos años, buscando continuamente significados más útiles.

Sin embargo, he terminado la historia con la escala Srole, porque ilustra otro punto importante. Dejar que la conceptualización y operacionalización quede con un final abierto no produce necesariamente anarquía y caos como se podría esperar. Emerge el orden y hay varios elementos en este orden. Primero, aunque usted pudiera definir la anomía de la manera que eligiese -por ejemplo, en términos de número de calzado- probablemente no la defina de forma muy diferente a las imágenes mentales de otras personas. Si realmente utilizara una definición fuera de lugar, la gente probablemente le ignoraría.

Segundo, cuando los investigadores descubren la *utilidad* de una conceptualización y operacionalización particular de un concepto, es más probable que lo adopten y aparezcan definiciones estandarizadas de los conceptos. Además de la escala Srole, se

incluyen ejemplos de tests CI y múltiples mediciones demográficas y económicas desarrolladas por el Departamento del Censo. Utilizar tales medidas establecidas tiene dos ventajas: han estado ampliamente comprobadas y descontaminadas y pueden compararse los estudios que emplean las mismas escalas. Si usted y yo realizamos estudios separados de dos grupos diferentes, y si cada uno de nosotros emplea la escala Srole, podríamos comparar nuestros dos grupos con respecto a la anomía.

Por tanto, los científicos sociales pueden medir cualquier cosa que sea real y podemos incluso hacer un buen trabajo midiendo cosas que no lo son. Garantizando que aunque algunos conceptos como el estatus socioeconómico, el prejuicio, la compasión y la anomía no sean finalmente reales, ahora ya hemos visto que los científicos sociales son capaces de crear orden cuando los manejan. Sin embargo, es un orden basado en la utilidad y no en una verdad absoluta.

El resto del capítulo está dedicado a algunas de las consideraciones y alternativas incluidas en la creación de definiciones y medidas útiles. Primero, vamos a ver la relación entre las definiciones y los objetivos de la investigación; más tarde, el capítulo termina con un examen de algunos criterios utilizados para determinar la calidad de las medidas que creemos.

DEFINICIONES Y OBJETIVOS DE LA INVESTIGACIÓN

Como recordará del Capítulo 4, los dos objetivos generales de la investigación son la *descripción* y la *explicación*.

LOS ORÍGENES DE LA ANOMÍA

por Leo Srole

Centro de Geriatria y Gerontología, Universidad de Columbia.

Mi interés por la anomía comenzó cuando leí *El Suicidio* de Durkheim, siendo estudiante de Harvard. Más tarde, como estudiante licenciado en Chicago, fui alumno de dos antropólogos seguidores de Durkheim: William Lloyd Warner y Alfred Radcliffe-Brown. Radcliffe-Brown había mantenido una larga correspondencia con Durkheim, haciéndome "descendiente" del gran sociólogo francés.

Para mí, el temprano impacto del trabajo de Durkheim sobre el suicidio fue desigual pero permanente. Por un lado, yo tenía serias reservas sobre su tenaz, ingenioso y a menudo delicado esfuerzo para forzar los datos burocráticos y toscos sobre las tasas de suicidio y hacer que se ajusten con su determinismo sociológico unidireccional. Por otro lado, yo estaba removido por la obsesiva preocupación con la fuerza moral de los lazos interpersonales que nos atan al tiempo, lugares y pasado y también con sus descubrimientos sobre las consecuencias letales que pueden seguir a la pérdida y deterioro de esos lazos.

Mi interés por la anomía recibió una sacudida testimonial a finales de la segunda guerra mundial, cuando serví en el Socorro de las Naciones Unidas y en la Administración para la Rehabilitación, ayudando a reconstruir una

Europa devastada por la guerra. En el campo de concentración nazi de Dachau, yo vi con mis propios ojos la profunda deshumanización que las fuerzas macrosociales, como las que atraparon a Durkheim, podían producir en individuos como Hitler, Eichmann, y otros que les servían en cualquiera de las fábricas mortales nazis.

Al volver de esta misión, sentí que nos habíamos retrasado demasiado en llegar a comprender la dinámica que subyace a la desintegración de los lazos sociales. Necesitábamos trabajar exhaustivamente, desestimando la proliferación de la teoría del macro-nivel en favor de un encuentro exploratorio directo con los individuos, utilizando una nueva metodología de investigación. También sentí que esta investigación debiera centrarse en un espectro más amplio de las patologías conductuales que el suicidio.

Mis investigaciones iniciales requirieron esfuerzos diversos. Por ejemplo, en 1950, pude entrevistar a una muestra de 401 conductores de autobús en Springfield. Cuatro años después el estudio de Salud Mental del centro de Manhattan, me permitió llegar a una gran cantidad de población. Estos y otros estudios de campo me dieron la perspectiva para ampliar y refinar mis mediciones de esta cualidad en individuos que reflejaban la cualidad macro-social que Durkheim había denominado "*anomie*".

Aunque empecé usando el término de Durkheim en mi propio trabajo, pronto decidí que era necesario limitar la utilización de ese concepto a su significado macro-social y separarlo de sus manifestaciones individuales. Para este último propósito, empleé el análogo pero hasta ahora obsoleto término griego, "*anomia*".

Publiqué por primera vez el constructo anomía en un artículo de 1956 en la *American Sociological Review**, describiendo maneras de operacionalizarlo y presentado los resultados de su inicial investigación de campo aplicada. Para 1982, el Science Citation Index y el Social Science Citation Index tenían una lista de 400 publicaciones en revistas de ciencias políticas, psicología, trabajo social y sociología en Estados Unidos y en el extranjero, que citaban el uso de estos instrumentos o hallazgos del artículo, garantizando que el Instituto Americano de Información Científica lo designa como una "cita clásica".

*Leo Srole, "Social Integration and Certain Corollaries. M Exploratory Study". *American Sociological Review*. Vol. 21, 709-16, 1956.

La distinción entre ambas tiene implicaciones importantes en el proceso de definición y medida. Si se ha formado la opinión de que la descripción es una tarea más sencilla que la explicación, se sorprenderá al conocer que las definiciones son más problemáticas para la investigación descriptiva que para la investigación explicativa. Este punto se discutirá más en profundidad en la Parte 4, pero es importante que tenga una comprensión básica de por qué es así, antes de que abordemos otros aspectos de la medición.

La importancia de las definiciones en la investigación descriptiva debe quedar clara. Si nuestra tarea es describir e informar de la tasa de desempleo de una ciudad,

nuestra definición de *estar desempleado* es crítica. Esa definición va a depender de nuestra definición de otro término: la *fuerza de trabajo*. Parece bastante absurdo considerar a un niño de tres años como un desempleado, porque no se le considera un miembro de la fuerza de trabajo.

Sin embargo, este acuerdo por sí solo no nos daría una definición satisfactoria, porque consideraríamos desempleados a personas que están estudiando en la universidad, a los retirados, incapaces y amo/as de casa. Lo más conveniente es conseguir un censo, definiendo la fuerza de trabajo como "todas las personas de 14 ó más años que están trabajando, buscando trabajo o esperando que les vuelvan a llamar de un empleo del que habían sido despedidos. Entonces, las personas desempleadas serían aquellos miembros de la fuerza de trabajo que no están empleados. Si un estudiante, amo/a de casa o persona retirada no está buscando trabajo, no se le incluiría dentro de la fuerza de trabajo.

Pero, ¿qué significa buscar trabajo?, ¿debe registrarse una persona en la oficina de empleo o ir pidiéndolo de puerta en puerta?, ¿o sería suficiente con querer un trabajo o estar abierto a una oferta de empleo? Convencionalmente, "buscar trabajo" se define operacionalmente como responder sí a la pregunta: "¿ha estado buscando empleo durante los últimos siete días?" (siete días es el período que se suele tomar, pero en función de los objetivos de la investigación se puede alargar o acortar ese período).

Me he parado a analizar estas consideraciones con detalle para que usted pueda darse cuenta de que la conclusión de un estudio descriptivo sobre la tasa de desempleo, por ejemplo, depende directamente de cómo se resuelve cada hecho. Aumentando el período de tiempo durante el cual contamos las personas que están buscando trabajo, se produciría el efecto de añadir más desempleados a la fuerza de trabajo y por tanto, aumentaría la tasa de desempleo. Si seguimos otro criterio y hablamos de la fuerza de trabajo *civil* y la tasa de desempleo *civil*, estamos excluyendo al personal militar; esto también aumenta la tasa de desempleo, porque el personal militar estaría trabajando - por *definición*-.

Por tanto, la afirmación descriptiva de que la tasa de desempleo en una ciudad es del 3 por ciento, o del 9 por ciento, depende directamente de las definiciones operativas que usemos. Si este ejemplo resulta claro es porque hay un número de acuerdos ya aceptados en relación a la fuerza de trabajo y al desempleo. Considerarnos lo difícil que sería ponerse de acuerdo con respecto a las definiciones necesarias para hacer la afirmación descriptiva "el 15 por ciento de los estudiantes son políticamente conservadores". Este porcentaje, como la tasa de desempleo, dependería directamente de su definición de lo que debe medirse. Se partiría de una definición diferente para llegar a la conclusión: "el 5 por ciento del cuerpo de estudiantes es políticamente conservador".

Irónicamente, las definiciones son menos problemáticas en el caso de la investigación explicativa. Supongamos que estamos interesados en explicar el conservadurismo político. ¿Por qué hay personas conservadoras y otras no? Concretando más, supongamos que nos interesa si las personas mayores son en general más conservadoras que las jóvenes. Si usted y yo tenemos 25 definiciones operativas diferentes de *conservador*, ¿cómo haremos si no nos ponemos de acuerdo en cuál es la mejor? Como ya hemos visto, esto no es necesariamente un obstáculo insalvable para nuestra investigación. Por ejemplo, supongamos que hemos encontrado que los mayores son más conservadores que los jóvenes en función de las *¡25 definiciones!* (Recuerde la

discusión anterior sobre -la compasión en hombres y mujeres). Supongamos que hemos averiguado que las personas mayores son más conservadoras que las jóvenes en *todas* las definiciones de conservadurismo que consideramos razonables. No importaría cuál fuera nuestra definición, llegaríamos a la conclusión de que las personas mayores son en general más conservadoras que las jóvenes, incluso aunque no nos pusiéramos de acuerdo en qué es realmente ser conservador.

En la práctica, la investigación explicativa rara vez resulta tan ambigua como en el ejemplo. Al contrario, el patrón general es bastante común en la investigación actual. *Hay* patrones sistemáticos de las relaciones en la vida social humana, que producen hallazgos sistemáticos en la investigación. Sin embargo, lo importante es que tal sistematización no aparece en una situación descriptiva. El cambio en las definiciones casi resulta inevitable en las diferentes conclusiones descriptivas.

El apartado "La importancia de los nombres de las variables" explora este hecho en conexión con la variable "participación ciudadana".

CRITERIOS PARA MEDIR CON CALIDAD

Hemos ido avanzando en este capítulo. Comenzamos con la simple afirmación de que los científicos sociales pueden medir cualquier cosa que exista. Después, descubrimos que la mayoría de las cosas que quisiéramos medir y estudiar no existen en la realidad. También aprendimos que de todas maneras es posible medirlas. Pero quisiera concluir el capítulo con un comentario sobre algunos de los baremos con los que juzgamos nuestro éxito o fracaso relativo, cuando medimos cosas -incluso cosas que no existen.

Para empezar, las mediciones pueden tener diferentes grados de precisión, que representan la fuerza de distinciones hecha entre los atributos que componen una variable. La descripción de una mujer de "43 años" es más precisa que decir que "se encuentra en los cuarenta". Decir "un metro y cuatro centímetros" es una descripción más precisa que "algo más de un metro".

Como norma general, las mediciones precisas son superiores a las imprecisas, por sentido común. No hay ninguna condición bajo la cual las mediciones imprecisas serían intrínsecamente superiores a las precisas. Sin embargo, la precisión no siempre es necesaria o deseable. Si el objetivo de su investigación es tal que saber que una mujer se encuentra en los cuarenta es suficiente, entonces cualquier esfuerzo adicional para conocer su edad precisa es una pérdida de tiempo. Entonces, la operacionalización de los conceptos debe guiarse en parte por una comprensión del grado de precisión que se requiere. Si sus necesidades no están claras, sea lo más preciso posible. Pero no hay que confundir precisión con *exactitud*. Describir a alguien como "nacido en Stowe, Vermont" es más preciso que "nacido en Nueva Inglaterra", pero supongamos que la persona en cuestión nació realmente en Boston. En este ejemplo, la descripción menos precisa habría sido la más exacta, el mejor reflejo de la realidad.

Precisión y exactitud obviamente son cualidades importantes de las mediciones en investigación y probablemente no necesiten más explicación. Sin embargo, cuando los científicos sociales construyen y evalúan las mediciones, prestan especial atención a dos consideraciones técnicas: fiabilidad y validez.

Fiabilidad

La *fiabilidad* consiste en si una técnica en particular, aplicada repetidamente al mismo objeto, conseguiría siempre el mismo resultado. Supongamos, por ejemplo, que le pido que calcule mi peso. Usted me mira detenidamente y calcula que peso 75 kilos. (Gracias). Ahora supongamos que le pido que calcule el peso de otras 30 ó 40 personas y mientras se encuentra enfrascado en esta tarea, yo me cuelo en la cola de nuevo pero disfrazado. Cuando me toca otra vez el turno, usted calcula 80 kilos. ¡Vaya! Este pequeño ejercicio le demostraría que ponerle a calcular el peso de la gente no es una técnica muy fiable.

Sin embargo, supongamos que le hubiera prestado mi báscula de baño para que pesase a la gente. No importa lo bien disfrazado que fuese porque presumiblemente usted hubiera anunciado el mismo peso las dos veces, lo cual indica que la escala que utiliza es una medida para pesar más fiable que adivinarlo.

LA IMPORTANCIA DE LOS NOMBRES DE LAS VARIABLES

por Patricia Fisher

Escuela de Graduados en Planificación. Universidad de Tennessee.

La operacionalización es una de esas cosas que es más fácil decir que hacer. Es bastante sencillo explicar a alguien el objetivo y la importancia de las definiciones operativas de las variables e incluso describir cómo se realiza un operacionalización típica. Sin embargo; hasta que no se intenta operacionalizar una variable compleja, no se pueden apreciar las dificultades que se presentan. De importancia considerable en la operacionalización es el nombre particular que usted haya elegido para la variable. Veamos un ejemplo extraído del área de planificación urbanística.

Una variable interesante para los planificadores es la participación ciudadana. Los planificadores están convencidos de que la participación de los ciudadanos en el proceso de planificación es importante para alcanzar el éxito en su implementación. La participación ciudadana es una ayuda para que los planificadores comprendan las necesidades reales y percibidas de una comunidad, y esta implicación de los ciudadanos tiende a aumentar la cooperación y a apoyar los esfuerzos de planificación. Aunque podrían escribirse muchas definiciones conceptuales diferentes de cada planificador, apenas habría desacuerdo o malos entendidos respecto a lo que denominamos participación ciudadana. El nombre de la variable parece adecuado.

Sin embargo, si pedimos a diferentes planificadores que nos ofrezcan medidas operativas muy simples para la participación ciudadana, es probable que haya varias respuestas que generen confusión. Un planificador podría llevar la cuenta de las asistencias de ciudadanos particulares a la comisión ciudadana y a otros encuentros del gobierno local; un segundo podría registrar el número de asistencias a un encuentro del gobierno local, las cartas y llamadas telefónicas recibidas por el alcalde y otros cargos públicos, y los encuentros que se siguen con un interés especial durante un determinado

período de tiempo. Como haría un hábil investigador, podemos ver rápidamente que cada planificador estaría midiendo (de manera un poco simplista) una dimensión diferente de la participación ciudadana: la amplitud de la participación ciudadana, temas que les preocupan y forma de participación. Por lo tanto, el nombre original de nuestra variable, la participación ciudadana, que era bastante satisfactorio desde el punto de vista conceptual, se muestra inadecuado para la operacionalización.

Es importante en una investigación, que se nombre a las variables de manera precisa y exacta, como resultado de una buena operacionalización. A menudo los nombres de las variables son el resultado de un proceso interactivo de formación de una definición conceptual, después una definición operativa y después se renombra el concepto para manejar mejor lo que vamos a medir. Este proceso circular continúa (el ejemplo anterior ilustra sólo una de las vueltas), y se va refinando gradualmente el nombre de la variable y su medición, hasta que se consiga un resultado satisfactorio. Algunas veces, el concepto de la variable con el que se termina es un poco diferente del original con el que ha empezado, pero por lo menos sabe que está midiendo lo que habla, ¡a lo mejor sólo porque está hablando de lo que mide!

Pero la fiabilidad no asegura la exactitud más allá de lo que la asegura la precisión. Suponga que he trucado mi báscula de baño para que marque 3 kilos menos con el fin de sentirme mejor. Aunque usted informaría (con fiabilidad) del mismo peso las dos veces, estaría equivocado. Este nuevo elemento se denomina *sesgo* y se tratará con más detalle en el Capítulo 8, que habla sobre la muestra. Por tanto, advierta que la fiabilidad no asegura la exactitud.

Supongamos que nos interesa estudiar el estado de ánimo de los trabajadores de dos clases diferentes de fábricas. En un grupo de ellas, los empleados realizan trabajos muy especializados, reflejando una extrema división del trabajo. Cada trabajador contribuye a una pequeña parte del proceso total, que se desarrolla en una larga cadena de montaje. En el otro grupo de fábricas, cada trabajador desarrolla muchas tareas y grupos pequeños de trabajadores completan el proceso total.

¿Cómo deberíamos medir el estado de ánimo? Siguiendo una estrategia, podríamos pasar más tiempo observando a los trabajadores de cada fábrica, fijándonos en si se hacen bromas unos a otros, si sonríen y ríen mucho, etc. Podríamos preguntarles cómo se sienten con su trabajo e incluso si creen que preferirían su trabajo actual a otro que se les ofreciera. Comparando lo que observamos en las diferentes fábricas, podríamos llegar a una conclusión sobre qué proceso de montaje produce mejor estado de ánimo.

Ahora vamos a ver algunos de los posibles problemas de fiabilidad inherentes a este método. Primero, cómo nos sentimos usted y yo cuando hacemos una observación parecida. Podemos malinterpretar lo que vemos. Podemos ver a los trabajadores gritándose y pensar que mantienen una discusión. O puede ser que les pillemos en un día que no han ido a trabajar. Si observamos al mismo grupo de trabajadores varios días peleándose, podríamos llegar a diferentes evaluaciones cada día. Si varios observadores

evaluaran la misma conducta, por otro lado, también podrían llegar a diferentes conclusiones sobre el estado de ánimo de los trabajadores.

Vamos a conocer otra estrategia para evaluar el estado de ánimo. Supongamos que vamos a comprobar una compañía de grabación para ver cuántas quejas ha tenido con los sindicatos durante un período establecido de tiempo. Presumiblemente, esto sería un indicador del estado de ánimo: cuantas más quejas, peor estado de ánimo. Esta estrategia de medición parece que es más fiable: Podríamos contar las quejas una y otra vez y deberíamos obtener siempre el mismo número.

Si se descubre a sí mismo diciendo: "¡Un momento!", con respecto a la segunda estrategia de medición, se estará preocupando de la validez, no de la fiabilidad. Vamos a terminar con la fiabilidad y después nos ocuparemos de la validez.

Los problemas de fiabilidad sesgan de muchas formas la investigación social. Los investigadores saben desde hace mucho tiempo que diferentes entrevistadores obtienen diferentes respuestas de los encuestados, como resultado de sus propias actitudes y comportamientos. Si tenemos que llevar a cabo un estudio del posicionamiento de los editoriales de los periódicos sobre algún acontecimiento público, podríamos crear un equipo de codificadores para que se tomen el trabajo de leer cientos de editoriales y los clasifiquen en función de su posicionamiento con respecto al acontecimiento. Diferentes codificadores codificarían el mismo editorial de manera diferente. También podríamos querer clasificar cientos de ocupaciones específicas en función de algunos esquemas de codificación establecidos, como un grupo de categorías creadas por el Departamento de Trabajo o por el Departamento del Censo. Usted y yo no codificaríamos todas esas ocupaciones en las mismas categorías.

Cada uno de estos ejemplos ilustra problemas de fiabilidad. Surgen problemas similares cada vez que pedimos a las personas que nos den información de sí mismas. Algunas veces preguntamos cosas que la gente no sabe responder: (¿Cuántas veces ha ido a la iglesia?). Algunas veces hacemos preguntas que les son totalmente irrelevantes: (¿Está usted satisfecho con la actual relación entre China y Albania?). Y otras veces hacemos preguntas que son tan complicadas que una persona que tenía una opinión clara sobre el tema, podría llegar a una interpretación diferente de la pregunta cuando se la hacemos por segunda vez.

¿Cómo puede crear medidas fiables? Existen algunas técnicas. En primer lugar, al pedir información a la gente -si su diseño de investigación se lo exige- formule preguntas sólo sobre cosas importantes para ellos, y haga la pregunta con claridad. El peligro en estos casos es que las personas le darán respuestas fiables o no. Habrá quien le diga cómo se siente ante la relación de China con Albania, incluso aunque no tenga ni la más remota idea de en qué consiste esa relación.

Incluso cuando los encuestados pueden responder a las preguntas, podría ocurrir que las respuestas careciesen de fiabilidad. Consideremos el dilema de los periodistas políticos la víspera de la Convención Democrática de 1984:

Al oír lo que nos dice la NBC, Mondale tiene 415 votos potenciales; Gary Hart, 236 y Jesse Jackson 36, mientras que 165 están indecisos.

Según la CBS, Mondale tiene 752; Hart, 45, Jackson, 88 y 448 están indecisos.

Sin embargo, ABC, contabiliza 755 para Mondale; 444 para Hart; 72 para Jackson; 1 para John Glenn (que se ha retirado); y 355 indecisos.

La Associated Press y la United Press International tienen la misma variación en los votos totales. No hay dos fuentes que se pongan de acuerdo (*The Chronicle* utiliza los totales de la Associated Press, que muestra a Mondale con 692 delegados; Hart, 422; Jackson, 76, y 160 indecisos).

(*San Francisco Chronicle*, 30/3/84, p. 12)

Una razón importante para las discrepancias sugeridas en el artículo, era que "muchas de estas personas han anunciado a quién van a apoyar, y otros habían sido escrutados en función de los votos contabilizados. Pero no todo el mundo da la misma respuesta a la misma pregunta".

El problema de la fiabilidad es básico en la medición dentro de las ciencias sociales y los investigadores han desarrollado algunas técnicas para manejarlo.

Método del Test-Retest. A veces es apropiado realizar la misma indicación más de una vez. Si no hay ninguna razón para esperar que la información sufra un cambio, entonces debiera esperar la misma respuesta ambas veces. Sin embargo, si las respuestas varían, eso puede indicar que el método de medición no es fiable hasta el grado de esa variación. Vamos a poner un ejemplo.

En su investigación sobre Valoración de los Riesgos de Salud (VRS), una parte de la medicina preventiva, Jeffrey Sacks y sus colegas (1980) quisieron determinar los riesgos asociados con varios factores de fondo y estilos de vida, para facilitar que los médicos aconsejen a sus pacientes de manera apropiada. Conociendo las situaciones vitales de los pacientes, los médicos podrían advertirles de los peligros sobre su salud y cómo mejorarla. Por supuesto, este objetivo depende en gran parte de la exactitud de la información recogida en el estudio sobre cada sujeto.

Para comprobar la fiabilidad de su información, Sacks y sus colegas pidieron a 207 sujetos que completaran, como línea base, un cuestionario con preguntas sobre sus características y conducta. Tres meses después, se les pasó un cuestionario de seguimiento a los mismos sujetos con la misma información y se compararon los resultados de los dos informes. Solo el 15 por ciento de los sujetos dieron la misma información en ambos estudios.

Según Sacks (1980: 730): "Casi el 10 por ciento de los sujetos dijeron que tenían una altura diferente en el examen de seguimiento. Tres sujetos cambiaron la edad de sus padres. Un padre aparentemente creció 20 años cronológicos en tres meses. Cinco ex-fumadores y ex-bebedores tuvieron dificultades en recordar su patrón previo de consumo".

Algunos sujetos habían borrado todo rastro anterior de taquicardia, diabetes, enfisema y pensamientos de suicidio. La madre de un sujeto que estaba muerta en el primer cuestionario, aparentemente estaba viva y sana en el segundo. Una sujeto que había perdido un ovario en el primer estudio, lo tenía en el segundo. En otro caso, el ovario presente en el primero, se perdió en el segundo - ¡pero desde hacía diez años!-. Un sujeto dijo tener 55 años en el primer estudio y 50 años tres meses después. Habría que preguntarse si los consejeros médicos podían haber tenido más éxito anotando personalmente los datos de sus pacientes que fiándose de sus memorias. Por eso, el método de recogida de datos no era especialmente fiable.

Método de las Dos Mitades. Como norma general, siempre es una buena idea tomar más de una medida de cualquier concepto social complejo o sutil como el prejuicio, alienación o la clase social. Este procedimiento también exige comprobar la fiabilidad. Imaginemos que ha creado un cuestionario que contiene diez ítems que usted piensa que miden el prejuicio contra las mujeres. Utilizando *la técnica de las dos mitades*, dividiría al azar (ver Capítulo 8) esos diez ítems en dos grupos de cinco. Como ya vimos en la discusión de Lazarsfeld “indicadores intercambiables”, cada grupo debiera proporcionar una buena medida del prejuicio contra las mujeres, y debiera corresponder a lo que se mide en el estudio. Si los dos grupos de ítems miden a las personas de modo diferente, otra vez tenemos un problema en la fiabilidad de la medida de dicha variable.

Utilizar Medidas Establecidas. Otra manera de enfrentarse al problema de la fiabilidad cuando se recoge información de personas, es utilizar medidas que hayan probado su fiabilidad en investigaciones previas. Si usted quiere medir la anomia, por ejemplo, podría seguir la línea de Srole.

Es importante reconocer que el utilizar muchas veces la misma medida no garantiza su fiabilidad. Por ejemplo, el Test de Aptitudes Escolares y el inventario de Personalidad Multifásica de Minnesota (MMPI), se han aceptado como medidas establecidas y reconocidas en sus respectivos dominios durante décadas. Recientemente, han necesitado una revisión que refleje los cambios sociales.

Fiabilidad del Investigador. También es posible que la falta de fiabilidad en las medidas se deba al investigador que la realiza: por ejemplo, a los entrevistadores y codificadores. Hay varias soluciones. Para protegerse de la falta de fiabilidad del entrevistador, suele ser habitual tener un supervisor que llame por teléfono a una submuestra de los encuestados y verifique parte de la información seleccionada.

También se puede hacer una réplica. Si usted está preocupado de si los editoriales o solicitudes de empleo de los periódicos se han clasificado con fiabilidad, ¿por qué no pide que varios codificadores los clasifiquen por separado? Cuando haya un desacuerdo entre ellos, se evalúa con más cuidado y se resuelve.

Finalmente, la claridad, especificidad, entrenamiento y práctica, evitarán en gran medida la falta de fiabilidad y el sufrimiento. Si dedicamos más tiempo a llegar a un acuerdo sobre cómo vamos a evaluar las opiniones editoriales sobre un tema -discutiendo las distintas posiciones que podrían tomarse y leyendo varias juntas-, probablemente seríamos capaces de hacer un buen trabajo de clasificación.

La fiabilidad de las medidas es un aspecto fundamental en la investigación social y volveremos a él más de una vez en los capítulos siguientes.

Sin embargo, por ahora, recordemos que incluso la fiabilidad total no asegura que nuestras medidas midan lo que pensamos que miden. Pasemos ahora a la validez.

Validez

El término *validez*, en su acepción convencional, se refiere hasta dónde una medida empírica refleja adecuadamente el *significado real* del concepto que estamos considerando. Pero ya he dicho que los conceptos no tienen significados reales. Entonces,

¿cómo podemos decir si una medida en particular refleja adecuadamente el significado del concepto? Por supuesto, no podemos. Además, ya he apuntado algunas de las maneras en las que los investigadores manejan este hecho.

Primero, existe la denominada validez aparente. Medidas empíricas particulares pueden o no concordar con nuestros acuerdos e imágenes mentales asociados a un concepto particular. Usted y yo podríamos discutir sobre si es adecuado medir el estado de ánimo de un trabajador contando el número de quejas presentadas al sindicato, pero seguramente estaríamos de acuerdo en que el número de quejas tiene algo que ver con su estado anímico. Si yo fuera a sugerir que midiéramos el estado de ánimo averiguando cuántos libros sacan los trabajadores de la biblioteca durante sus ratos libres, usted indudablemente presentaría una objeción mucho más seria: esa medida no tendría ninguna validez aparente.

Segundo, ya he indicado muchos de los acuerdos concretos que los investigadores han alcanzado con respecto a algunos conceptos. El Departamento del Censo, por ejemplo, ha elaborado definiciones operativas de conceptos como familia, cabeza de familia, estatus laboral, para que se trabaje mejor con la validez en los estudios que utilizan estos conceptos.

Edward Carmines y Richard Zeller (1979) citan tres tipos de validez: *validez referida al criterio*, *validez de constructo* y *validez de contenido*.

La validez referida al criterio también se le llama algunas veces *validez predictiva* y está basada en algún criterio externo. Por ejemplo, la validez de una prueba de capacidad para los estudios se demuestra por su habilidad para predecir el éxito de los estudiantes en la escuela. La validez de un test de conducir escrito se determina, en este sentido, por la relación entre las puntuaciones que los individuos obtienen en el test y cómo conducen. En estos ejemplos, el éxito escolar y la habilidad para conducir son los criterios. En general, la conducta puede servir como una medida de la validez del criterio para la mayoría de las medidas actitudinales que realizamos en investigación social -por ejemplo, ¿realmente las personas "con prejuicios" discriminan a las minorías?- aunque la relación entre actitudes y conducta es también una cuestión de estudio importante en sí misma.

Algunas veces es difícil encontrar criterios conductuales que puedan tomarse para validar las mediciones tan directamente como en los ejemplos anteriores. Sin embargo, en esos ejemplos a menudo podemos aproximarnos a tales criterios considerando cómo la variable en cuestión debiera, en teoría, relacionarse con otras variables. La validez de constructo está basada en las relaciones lógicas entre las variables. Por ejemplo, supongamos que usted está interesado en estudiar la "satisfacción marital", sus causas y consecuencias. Como parte de su investigación desarrollará una medida de la satisfacción marital y querrá evaluar su validez.

Además de desarrollar su medida, también tendrá que desarrollar algunas expectativas teóricas sobre la manera en que la variable satisfacción marital se relaciona con otras variables. Por ejemplo, podría concluir razonablemente diciendo que será menos probable que los maridos y esposas satisfechos engañen a su cónyuge que los insatisfechos. Si usted mide la satisfacción marital en relación a la fidelidad marital, eso constituye una evidencia de su validez de constructo. Sin embargo, si los matrimonios "satisfechos" engañaran a sus cónyuges tanto como los "insatisfechos", se cuestionaría la validez de su medición.

Así pues, las pruebas de validez de constructo pueden ofrecer evidencia de si su medición solapa o no la cualidad que quiere medir, sin tener una prueba definitiva. Como ya he dicho, las pruebas de validez de constructo son menos precisas que las de validez de criterio, sin embargo, existe desacuerdo sobre qué tipo de prueba es la más indicada para una variable en particular, en una situación dada. Es menos importante que diferencie estos dos tipos de validez, que entender la lógica de la validación que tienen en común: si hemos podido medir correctamente alguna variable, esas mediciones debieran relacionarse lógicamente con otras medidas.

Finalmente, la validez de contenido se refiere al grado en el que una medida cubre el rango de significados incluidos en el concepto. Por ejemplo, una prueba de habilidad matemática, según Carmines y Zeller, no puede limitarse a sumar, sino que también tendrá que restar, multiplicar, dividir, etc. O, si decimos que estamos midiendo el prejuicio *en general*, ¿nuestras medidas reflejarán el prejuicio contra los grupos étnicos y raciales, las minorías religiosas, las mujeres, los ancianos, etc.?

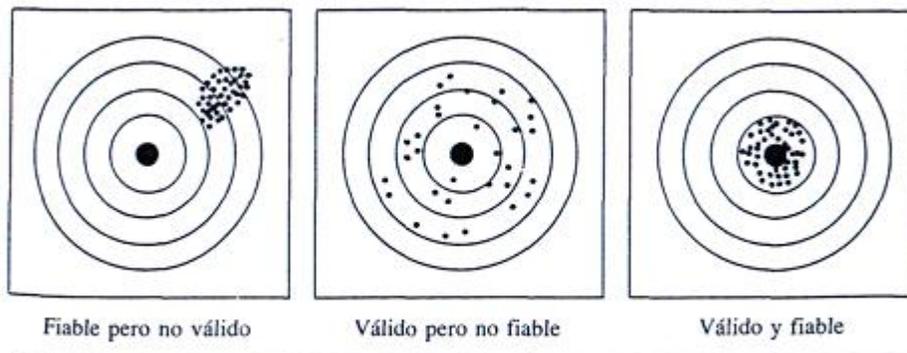


Figura 5.2. Analogía de la Validez y Fiabilidad. (Sugerido por un revisor anónimo)

La Figura 5.2 representa gráficamente la diferencia entre validez y fiabilidad. Imagínese una persona que dispara a una diana, verá que la fiabilidad parece un "patrón de coincidencia", independientemente del lugar de la diana en que se encuentre, puesto que la fiabilidad está en función de la consistencia. Por otro lado, la validez está en función de los disparos que acierten alrededor del centro de la diana. El tallo en la fiabilidad puede verse como un error debido al azar; el tallo de la validez es un error sistemático. Observe que no es muy, útil una medida que no sea fiable y válida.

Tensión entre Fiabilidad y Validez

Como complemento a estas explicaciones, quisiera decir brevemente que suele existir una cierta tensión entre fiabilidad y validez. A menudo nos vemos obligados a negociar entre las dos.

Si recuerda por un momento el ejemplo de cómo medir el estado de ánimo en diferentes fábricas, creo que reconocerá que la estrategia de sumergirse en la rutina diaria de una línea de montaje, observando cómo funciona y hablando con los trabajadores, parece proporcionar una medida más válida del estado de ánimo que contar el número de quejas. Parece obvio que podremos llegar a un conocimiento más

claro de si el estado de ánimo era mejor o peor, que contando el número de quejas presentadas al sindicato.

Sin embargo, como ya he señalado, la estrategia de contar las quejas sería más fiable. Esta situación refleja una tensión más general en las mediciones de la investigación. La mayoría de los conceptos realmente interesantes que queremos estudiar tienen muchos matices sutiles y es difícil especificar con precisión lo que significan. Algunas veces los investigadores dicen que estos conceptos tienen una gran "riqueza de significado". Se han escrito muchos libros y artículos sobre la anomía y todavía no se han agotado los aspectos interesantes de ese concepto.

La ciencia aún necesita ser específica para generar mediciones fiables. Por eso, muy a menudo la especificación de las definiciones y mediciones operativas fiables parece robar a tales conceptos su riqueza de significado. Lo que quiero decir es que, el estado de ánimo es mucho más que una ausencia de quejas a los sindicatos; la anomía es mucho más que cinco ítems creados por Leo Srole.

Se trata de un dilema inevitable y persistente para el investigador social, y se le puede hacer frente estando en guardia. Prepárese y manéjelo. Si no hay un acuerdo claro sobre cómo medir un concepto, médalo de diferentes maneras. Si 1 concepto tiene diferentes dimensiones, médalas todas. Y sobre todo, sepa que el concepto no tiene más significado que el que usted y yo queramos darle. La única justificación para dar a un concepto un significado particular es la utilidad. Mida los conceptos de forma que nos ayuden a entender el mundo que nos rodea.