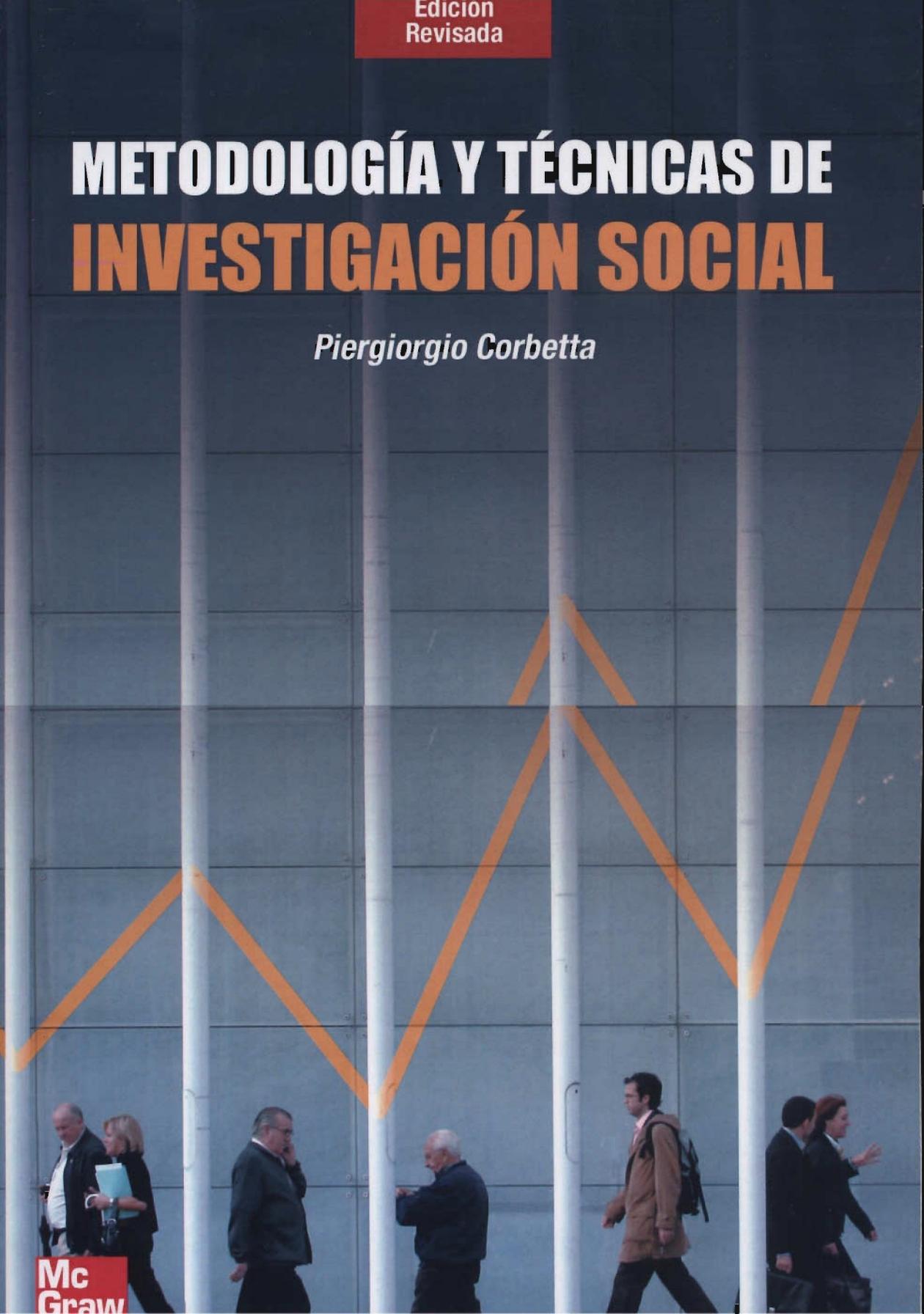


Edición
Revisada

METODOLOGÍA Y TÉCNICAS DE INVESTIGACIÓN SOCIAL

Piergiorgio Corbetta

Mc
Graw

The background of the book cover is a grey grid pattern. A blue line graph is overlaid on the grid, showing a fluctuating trend that generally moves upwards from left to right. At the bottom of the cover, there are silhouettes of several people walking in different directions, suggesting a busy, modern environment. The overall aesthetic is clean and professional.

LA TRADUCCIÓN EMPÍRICA DE LA TEORÍA

3.1. ESTRUCTURA «TIPO» DE LA INVESTIGACIÓN CUANTITATIVA	68
3.1.1. Las cinco fases del proceso de investigación	69
3.2. DE LA TEORÍA A LAS HIPÓTESIS	71
3.2.1. Teoría	71
3.2.2. Hipótesis	72
3.2.3. La transformación de la teoría en hipótesis	73
3.3. DE LOS CONCEPTOS A LAS VARIABLES	76
3.4. UNIDAD DE ANÁLISIS	79
3.4.1. Los distintos tipos de unidad de análisis	80
3.5. VARIABLES	81
3.6. VARIABLES NOMINALES, ORDINALES Y CARDINALES	84
3.6.1. Variables nominales	85
3.6.2. Variables ordinales	87
3.6.3. Variables cardinales	88
3.7. CONCEPTOS, INDICADORES E ÍNDICES	91
3.8. ERROR TOTAL	95
3.9. FIABILIDAD Y VALIDEZ	99
3.9.1. Fiabilidad	99
3.9.2. Validez	101
3.10. Síntesis del capítulo 3	102
3.11. LECTURAS COMPLEMENTARIAS	104

En este capítulo hablaremos de la operacionalización, es decir, la transformación de hipótesis teóricas en operaciones de una investigación empírica. En primer lugar explicaremos qué se entiende por teoría, hipótesis y conceptos. Después presentaremos la noción de «variable», que representa el concepto clave de la investigación social empírica, y analizaremos los diferentes tipos con los que se representa. Por último, abordaremos el problema general de los errores que se pueden cometer en el proceso de traducción de la teoría a la investigación empírica.

3.1. ESTRUCTURA «TIPO» DE LA INVESTIGACIÓN CUANTITATIVA

Para empezar conviene decir que *la investigación científica es un proceso de descubrimiento creativo que sigue un itinerario prefijado y unos procedimientos preestablecidos y consolidados dentro de la comunidad científica*. Este paralelismo entre la «creatividad» y los sucesivos «procedimientos preestablecidos» no debe parecer contradictorio. El atributo «creativo» es importante en esta definición, evoca las capacidades personales del investigador, su perspicacia e inteligencia intuitiva. Pero esto es sólo un aspecto del proceso de investigación científica.

Reichenbach introdujo una distinción entre la concepción de una idea nueva y la presentación, justificación, defensa y comprobación de la misma, momentos que denominó, respectivamente, *contexto del descubrimiento* y *contexto de la justificación*. Según Reichenbach, en el primer estadio no es posible establecer reglas ni procedimientos. «El acto del descubrimiento escapa al análisis lógico; no hay reglas lógicas con las que construir una “máquina descubridora” que desempeñe la función creativa del genio» [Reichenbach, 1951; 231]. Pero la labor del científico no consiste únicamente en producir teorías, sino también en comprobarlas. Y esta fase, tanto si aborda el examen lógico-formal de la teoría, como su congruencia con la realidad, debe seguir unas reglas muy precisas.

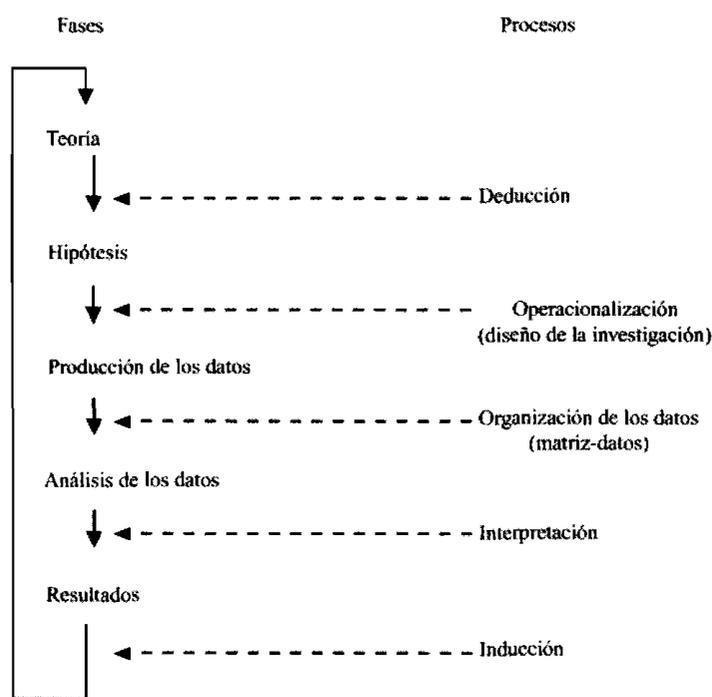
La primera regla fundamental de la investigación empírica es que debe desarrollarse dentro de un marco aceptado por la comunidad científica. La investigación científica, por contraste con el arte, es un proceso colectivo. «La ciencia es pública y no privada», escribió Merton. Este aspecto colectivo-público de la ciencia tiene una doble connotación. Por una parte, implica el *control*: «Los conceptos y procedimientos empleados, incluso por el sociólogo más intuitivo, deben estar estandarizados, y los resultados de sus intuiciones deben poder ser comprobados por otros investigadores» [Merton, 1968]. Por otra parte, implica la *acumulación*: «Si conseguí llegar más lejos que otros hombres, es porque me subí a hombros de gigantes», afirmó Isaac Newton, frase que Merton utilizó para construir un brillante ensayo sobre sociología de la ciencia [Merton, 1965]. Recordemos también que una de las definiciones más comunes de la ciencia es «acumulación sistemática de conocimientos».

El marco de referencia colectivo que debe tener presente el científico social al realizar su investigación, sobre todo en el enfoque cuantitativo, consta de dos elementos: la estructura lógica del proceso de investigación y la instrumentación técnica a utilizar. En el siguiente apartado trataremos el primero de estos elementos.

3.1.1. Las cinco fases del proceso de investigación

El *proceso* que sigue el investigador social para la realización de una investigación consiste en un recorrido cíclico que parte de la teoría, pasa por las fases de recopilación y análisis de los datos, y vuelve a la teoría. Este proceso se describe de manera más o menos similar en todos los manuales de investigación empírica, y para explicarlo podemos utilizar el esquema de la figura 3.1, donde distinguimos cinco fases y cinco procesos que conectan estas fases.

La primera fase es la de la *teoría*. La segunda es la de las *hipótesis*, y el paso entre las dos se produce mediante el proceso de *deducción*. La hipótesis representa una formulación parcial de la teoría, y se sitúa a un nivel inferior en cuanto a generalidad¹. La teoría es «general» y la hipótesis es «específica». Por ejemplo, de la proposición teórica general que afirma que «existe una correlación positiva entre la participación política y el prestigio social» pueden deducirse, entre otras, tres



FUENTE: Adaptado de Bryman [1988, 20].

Figura 3.1. Estructura de la investigación cuantitativa.

¹ Y algunas veces también en cuanto al ámbito de aplicación, es decir, el área geográfica o lapso de tiempo a los que se refiere la hipótesis.

hipótesis más específicas: que la participación electoral es superior en el caso de los varones, los adultos (en comparación con los jóvenes y los ancianos), y las personas con más éxito profesional. Al hacer esto, hemos limitado el concepto de participación política al de participación electoral, y el de prestigio social a los tres aspectos mencionados.

La tercera fase es la de *recopilación de datos*, y llegamos a ella mediante el proceso de *operacionalización*, que consiste en la transformación de las hipótesis en afirmaciones observables empíricamente. Este proceso es muy complejo y se puede dividir en dos partes. La primera consiste en la *operacionalización de los conceptos*, o sea, la transformación de los conceptos en variables, o entidades medibles. Por ejemplo, el concepto de éxito profesional se podrá operacionalizar mediante los ingresos, o la obtención de un empleo mejor que el de nuestros padres, etc. La segunda consiste en la *elección del instrumento* y los procedimientos para medir los conceptos. En el ejemplo de la participación política se podría trabajar con datos agregados, utilizando las estadísticas oficiales de voto y averiguando si la participación electoral varía con el género y la edad; también se podría recurrir a datos de encuestas de opinión pública, o incluso servirse de entrevistas en profundidad sobre una muestra de ciudadanos. La decisión sobre estos aspectos nos lleva a la definición del *diseño de la investigación*, un «plan de trabajo» que establece las diversas fases de la observación empírica (por ejemplo: se entrevistará a tantas personas, ubicadas en las siguientes áreas geográficas..., elegidas con el siguiente criterio..., utilizando el siguiente cuestionario..., etc.).

Después de la recopilación de material empírico, llegamos a la fase del *análisis de los datos*, que será precedida por la *organización y sistematización de los datos obtenidos*. Por lo general se da el nombre de *información* al material empírico bruto que aún no ha sido sistematizado, y el de *datos*, a ese mismo material organizado de tal forma que se pueda proceder a su análisis. Pensemos, por ejemplo, en un paquete de cuestionarios y varias cintas de entrevistas grabadas (información), y esos mismos cuestionarios o entrevistas transformados en secuencias de números, comparables entre sí y obtenidas atribuyendo un número a cada respuesta (datos).

El proceso de organización de los datos suele consistir en transformar la información en una matriz de números, la *matriz de datos*, llamada también matriz «casos por variables» ($C \times V$): ésta se obtendrá, por ejemplo, codificando de forma sistemática las respuestas a los cuestionarios, estandarizando las respuestas de las entrevistas abiertas, uniformando las fuentes estadísticas, aplicando una plantilla de análisis del contenido a los textos, etc. La matriz de datos obtenida representa el punto de partida para el análisis de los datos, que por lo general consiste en operaciones matemáticas realizadas con un programa de análisis estadístico.

La quinta fase consiste en la presentación de los *resultados*, mediante un proceso de *interpretación* del análisis estadístico realizado en la fase anterior. Por último, el investigador vuelve al punto inicial, es decir, a la *teoría*, mediante un proceso de *inducción*, que consiste en comparar los resultados empíricos con las hipótesis teóricas y el conjunto de la teoría, para confirmarla o reformularla.

Como puede verse, el esquema que hemos expuesto empieza con la teoría y

vuelve a ella, en un proceso cíclico que podríamos visualizar como una *espiral*, donde teoría e investigación están ligadas en el proceso de acumulación de conocimientos. Nótese que la sucesión descrita representa el proceso ideal para la investigación social de tipo cuantitativo, y que el esquema básico puede sufrir variaciones significativas en cada aplicación concreta del mismo. Por ejemplo, en relación con la teoría, el equilibrio entre deducción e inducción puede variar. Aunque el esquema básico prevé que la investigación no comience antes de que se haya realizado un estudio exhaustivo de la teoría y se hayan formulado hipótesis muy precisas, hay investigaciones que parten de un embrión de teoría, y ésta no se desarrolla ni toma cuerpo hasta la fase de análisis del material empírico. Otras veces se obvia toda la fase de la recopilación de datos, incluidos los procesos de operacionalización y organización de los mismos, porque se usan datos ya existentes para la investigación. Pero en cualquier caso el esquema de referencia sigue siendo el que aparece en la figura 3.1.

En este capítulo trataremos los procesos de deducción y operacionalización. En el apartado siguiente nos ocuparemos del primero. El resto del capítulo lo dedicaremos íntegramente al proceso de operacionalización de conceptos: la atención dedicada a este tema queda justificada por su complejidad y su relevancia en el proceso de investigación.

3.2. DE LA TEORÍA A LAS HIPÓTESIS

3.2.1. Teoría

Podemos definir una *teoría* como un conjunto de proposiciones relacionadas de manera orgánica, que se formulan con un alto grado de abstracción y generalización respecto a la realidad, se deducen de hechos empíricos que se producen con regularidad y de las que pueden inferirse predicciones empíricas. A continuación explicamos con más detalle los cuatro elementos de esta definición.

Conjunto de proposiciones: sistema coherente de afirmaciones que con frecuencia tiene estructura causal. *Abstracción y generalización*: la teoría va más allá de las expresiones empíricas concretas, en el aspecto conceptual y en el campo de aplicación. *Se deduce de hechos empíricos que se producen con regularidad*: la teoría surge de la constatación de repeticiones en la realidad observada, y se confirma con ellas. *Inferencia de predicciones empíricas*: de una teoría que surge de la observación de hechos empíricos que se repiten se pueden deducir hechos en otros contextos diferentes.

Ilustramos lo apuntado hasta aquí con un ejemplo clásico: la teoría sobre el suicidio de Durkheim. Como es sabido, Durkheim identifica tres «tipos ideales» de suicidio: el suicidio egoísta, el altruista y el anómico. En su teoría encontramos la siguiente proposición causal que representa el fundamento teórico del suicidio egoísta: «Cuanto más elevada sea la tasa de individualismo en un determinado grupo social, mayor será la tasa de suicidios en ese grupo». Por individualismo entendemos una situación social y cultural en la que la persona es plenamente libre y está sola para tomar decisiones en la vida; en el extremo

opuesto se encuentra la situación de cohesión social, en la que la actividad de una persona está sometida a controles sociales de manera permanente, y sus decisiones vienen determinadas en gran parte por normas colectivas.

La proposición teórica enunciada, que vincula el individualismo con el suicidio, es una *proposición causal*, que presenta las características siguientes: *abstracción*, puesto que los conceptos mencionados de individualismo, cohesión social y suicidio tienen significados abstractos respecto a la realidad observada; *generalización*, dado que la conexión que Durkheim propone se considera válida para una gran parte de la sociedad, cuando no para todas las sociedades; *se deduce de hechos empíricos que se repiten*, ya que Durkheim, como describe en su investigación, dedujo y comprobó empíricamente su teoría mediante el análisis de las fuentes estadísticas de su época; genera *predicciones empíricas*, porque permite prever, por ejemplo, que Irlanda tendría una tasa de suicidios inferior a la de Inglaterra.

3.2.2. Hipótesis

Una proposición teórica debe poder expresarse en hipótesis específicas. Una hipótesis es una proposición que implica una relación entre dos o más conceptos, situada en un nivel inferior de abstracción y generalidad con relación a la teoría y que permite una traducción de ésta en términos que se puedan someter a prueba empírica. Las dos características distintivas de la hipótesis son, por una parte, su menor abstracción (o mayor concreción) en términos conceptuales, y su menor generalidad (o mayor especificidad), en términos de extensión, con relación a la teoría. Y por otra parte, su provisionalidad: la hipótesis es una afirmación provisoria que se debe comprobar, derivada de la teoría pero que precisa una comprobación empírica para poder confirmarse.

De la teoría general de Durkheim pueden deducirse una serie de hipótesis específicas. Por ejemplo, es de esperar una tasa de suicidio más alta en las sociedades protestantes que en las católicas, por el hecho de que la religión protestante confía las decisiones a la libertad individual, mientras que la católica es impositiva y normativa; una tasa de suicidio más baja en las personas casadas y con hijos, por el conjunto de exigencias, vínculos, controles y deberes que conlleva la situación familiar; una tasa de suicidio más baja cuando una sociedad atraviesa un periodo de crisis política o está en guerra, por la consolidación de los lazos sociales entre ciudadanos expuestos al mismo peligro; y así sucesivamente.

Estas hipótesis se pueden *comprobar de manera empírica* mediante definiciones operativas adecuadas: por ejemplo, se definirá el concepto de «integración familiar» considerando que las personas casadas y con hijos tienen el grado máximo de integración, y las personas que viven solas el grado mínimo, con los grados intermedios necesarios, y se verá si en los datos existe una asociación entre esta variable y el suicidio. También se podrá constatar, como hizo Durkheim, que en igualdad de condiciones en distintas sociedades, existe una relación entre la tasa de suicidio y el tipo de religión predominante.

La validez de una teoría depende de su capacidad para transformarla en hipóte-

sis que se pueden comprobar de manera empírica. *El criterio de la verificación empírica es el criterio del carácter científico.* Si una teoría es imprecisa y confusa, difícilmente se podrá transformar en hipótesis comprobables, y sin una confirmación empírica, no superará el ámbito preteórico de las conjeturas. En las ciencias sociales el riesgo de que se formulen teorías vagas, demasiado confusas como para operacionalizarlas, es especialmente elevado. En este aspecto compartimos la máxima de que «es mejor una teoría errónea que una teoría imprecisa».

3.2.3. La transformación de la teoría en hipótesis

A continuación presentamos algunos ejemplos de transformación de una teoría en hipótesis concretas. El primer ejemplo se refiere a la teoría formulada por Inglehart a mediados de los años setenta sobre el cambio de valores en los países occidentales. Según esta teoría, los valores de las generaciones que crecieron en la posguerra serían distintos de los valores de las generaciones anteriores, y el cambio consistiría en el desplazamiento de la atención prestada a aspectos como la seguridad física y económica hacia otras cuestiones como la calidad de vida, la autorrealización, y la satisfacción intelectual y estética, lo que supone pasar de una orientación predominantemente materialista a otra que se ha definido como «postmaterialista». Las causas de este cambio serían varias. Aquí comentaremos sólo un aspecto, quizás el más importante: los cambios de las condiciones de vida durante los años de formación de los ciudadanos.

Las generaciones de la posguerra no han sufrido las crisis económicas que afectaron a las generaciones que crecieron entre las dos guerras mundiales y durante la Gran Depresión; además, en la posguerra las sociedades occidentales conocieron una prosperidad sin precedentes que permitió satisfacer las necesidades de subsistencia a prácticamente toda la población. Por otra parte, las nuevas generaciones no han vivido la experiencia de la guerra, que no obstante, sí afectó directamente a sus padres y abuelos. Según la teoría de las necesidades, el hombre empieza a poner en primer plano objetivos no materiales, sólo después de conseguir la seguridad física y económica.

De esta teoría general pueden derivarse hipótesis específicas. Primero tendremos que encontrar en todos los países occidentales una diferencia notable en las escalas de valores de jóvenes y ancianos, como consecuencia del hecho de que la orientación de valores que se constituye en los años de formación tiende a mantenerse durante toda la vida aunque las condiciones sociales y ambientales cambien.

La segunda hipótesis se refiere a la distinta magnitud de esta diferencia generacional en los distintos países. Si la teoría es válida, la diferencia debería ser más marcada en países donde ha sido mayor la variación de la calidad de vida entre los que crecieron antes y después de la Segunda Guerra Mundial. Por ejemplo, en Alemania esta diferencia debería ser máxima: durante la Primera Guerra Mundial los alemanes sufrieron escasez de recursos, una fuerte inflación, depresión económica, graves conflictos internos y la tragedia, en términos de destrucción del país y pérdidas de vidas humanas, de la Segunda Guerra Mundial. En cambio, en la pos-

guerra vivieron un periodo de bienestar económico y paz social notables. El caso de Inglaterra debería ser opuesto: antes de la Segunda Guerra Mundial era el país europeo más rico, el único de los países europeos que no vivió la guerra en su territorio ni sufrió la invasión extranjera. Sin embargo, en la posguerra experimentó un estancamiento económico persistente que provocó niveles de bienestar inferiores a los de muchos otros países europeos y sin duda por debajo de los que tenía en la primera mitad del siglo. Por tanto, en este caso se puede esperar una diferencia menor en la escala de valores entre generaciones.

La tercera hipótesis concierne también a la diferencia entre países, pero se refiere no tanto al distanciamiento generacional como a la cifra absoluta de quienes abrazan los valores postmaterialistas. Éstos deberían ser más numerosos cuanto más rico sea el país, y, por tanto estarían presentes sobre todo en Bélgica, Holanda y Suiza, y, en menor medida en España, Grecia e Italia [cfr. Inglehart, 1977].

Todas las hipótesis que hemos mencionado se pueden operacionalizar fácilmente y, por tanto, se pueden someter a su comprobación empírica.

En este punto debemos hacer una aclaración sobre la conexión entre teoría e hipótesis. En la práctica, la investigación no siempre se realiza en el orden descrito, primero la elaboración de la teoría y después la formulación de hipótesis que sirven de guía para el «diseño de la investigación». Con frecuencia sucede que las hipótesis se formulan *después* de la recopilación de datos y se comparan con estos datos a posteriori.

Tomemos como ejemplo una investigación sobre decisiones escolares individuales. El autor analiza la decisión de seguir o no estudiando después de la enseñanza secundaria y primaria, comparando tres concepciones distintas de la acción individual. En primer lugar, el punto de vista «estructuralista» sostiene «que los actos humanos dependen de vínculos externos que no dejan espacio para tomar decisiones realmente importantes; en segundo lugar, el punto de vista de la «no intencionalidad», que considera que el individuo está condicionado, aunque no tanto por factores estructurales externos como por factores psicológicos internos, la mayoría inconscientes; y, finalmente, el punto de vista de la «intencionalidad», que «supone que los individuos actúan [...] de acuerdo con ciertos propósitos, persiguiendo sus objetivos lo mejor que pueden» [Gambetta, 1987, 8, 16].

Como puede verse, se trata de una formulación clásica de la teoría de la acción individual. Lo que nos interesa es que, para dar respuestas empíricas a estas preguntas, los autores de los dos ejemplos utilizan los datos *ya existentes* de dos encuestas por muestreo realizadas con fines más bien generales (descripción de las orientaciones políticas y culturales de los jóvenes). Las encuestas contenían también preguntas sobre la intención de seguir estudiando, que se utilizaron para responder a las cuestiones antes mencionadas. El autor parte de tres hipótesis concretas de conducta, y analiza sus datos para ver cuál de ellas se cumple efectivamente en el caso empírico analizado. En este ejemplo, la elaboración teórica, aunque es posterior a la recopilación de datos, precede al análisis de los mismos.

Otras veces se recurre a la teoría *después* de analizar los datos para explicar un hecho anómalo o un resultado inesperado. Por ejemplo, en una investigación sobre abstencionismo electoral en Italia realizada en los años ochenta [Corbetta y Parisi,

1987], los autores plantearon dos hipótesis: que el abstencionismo electoral es un fenómeno de *apatía* política debido a la poca capacidad de los partidos para movilizar al electorado con sus aparatos organizativos, o bien un fenómeno de *protesta* por parte de los electores para mostrar a los partidos descontento y resentimiento.

En la primera hipótesis los autores pronosticaron una mayor abstención en las capas más periféricas de la sociedad, menos accesibles para la capacidad de movilización y comunicación de los partidos (el caso extremo serían las mujeres ancianas de las zonas rurales del sur de Italia, donde las cuatro variables de género, edad, ruralismo-urbanización y norte-sur tienen los valores más marginales). En la segunda hipótesis los autores pronosticaron un mayor abstencionismo entre los jóvenes varones de las grandes ciudades del norte, clases sociales centrales expuestas a los flujos de nuevas ideas, lo que mostraría un desacuerdo entre la parte más moderna del país y la política de épocas anteriores.

Sin embargo, los datos arrojaron unos resultados sorprendentes, contrarios a las dos hipótesis. Las categorías más abstencionistas resultaron ser las mujeres de las capitales de provincia. Un análisis detallado del perfil de abstencionistas reveló que se trataba en gran parte de mujeres que habían votado anteriormente a la Democracia Cristiana. Para explicar este resultado imprevisto, los autores recurrieron a la teoría de las «presiones cruzadas», según la cual «cuando el elector es sometido de manera simultánea a estímulos de igual intensidad pero opuestos, es probable que eluda el conflicto de la elección [...] y opte por la abstención» [*ibidem*, 59]. La debilidad de la religión en una sociedad secularizada y la distancia deliberada de la Iglesia del compromiso político, junto con la existencia de un nuevo sistema de valores canalizado sobre todo en los medios de comunicación de masas, creaban en el electorado femenino, hasta entonces sensible al mensaje de la Iglesia, una situación de punto muerto entre los valores antiguos y nuevos que generaba incertidumbre y, por tanto, abstención. Este proceso se iniciaba en las grandes ciudades, lugares de cambio por antonomasia.

En el caso citado los autores recurrieron a una teoría diferente, utilizada en otras ocasiones para interpretar la conducta electoral. Otras veces puede descubrirse una nueva teoría durante la fase empírica. Un ejemplo muy conocido es el de los experimentos realizados en los años veinte en las fábricas de la Western Electric de Hawthorne (cerca de Chicago). Los experimentos pretendían estudiar un problema muy simple: conocer cómo afectaba la variación de la iluminación al rendimiento de los trabajadores de un departamento concreto. En un principio se constató que el aumento de la iluminación hacía aumentar la productividad. Pero después se vio que, contra todo pronóstico, al volver a los niveles de iluminación iniciales, la productividad seguía siendo elevada.

Tampoco se confirmaron las expectativas de los investigadores cuando se dividió a las trabajadoras en un grupo experimental (para el que se hacía variar la iluminación) y un grupo de control (para el que la iluminación se mantenía constante): ambos grupos tenían un rendimiento superior a la media, sin variaciones perceptibles vinculadas al nivel de iluminación. Entonces un equipo de sociólogos de la organización dirigido por Elton Mayo estudió el caso y se fijó en otros factores no ligados estrictamente a la comodidad ambiental, empleando otras técnicas

como la observación participante y la entrevista. De este modo se descubrió que la variable determinante del rendimiento laboral no era la iluminación (ni otras variables de carácter técnico u organizativo, como la duración del descanso para el café, etc.), sino el hecho de que el grupo de trabajadoras se sintiera centro de atención y objeto de un experimento.

A partir de los resultados de estos experimentos, se empezó a perfilar una nueva corriente teórica dentro de la sociología laboral, la de las denominadas «relaciones humanas», que incluía entre las condiciones para un rendimiento laboral óptimo variables psicológicas como la cohesión del grupo de trabajo, la percepción por parte del individuo de la importancia que tiene en su grupo de trabajo y la utilidad de la labor que desempeña, etc.

Sobre la alteración de la secuencia teoría-hipótesis-observación hay que decir que en ocasiones la observación viene antes que las hipótesis por causas de fuerza mayor. Es el caso del «análisis secundario», que consiste en aplicar un segundo análisis a datos recopilados previamente por otros investigadores (la investigación de Gambetta es un ejemplo de ello, así como *Crime in The Making*, de Sampson y Laub, que comentamos en el capítulo 2).

Conviene señalar que, a menudo, la teoría no está tan bien definida como para admitir hipótesis claras y precisas. A veces el tema es nuevo, desconocido, y por tanto la investigación se mueve en un plano eminentemente descriptivo. Otras veces la observación no parte de una teoría específica, sino que pretende abarcar una amplia gama de problemas, para permitir sucesivos análisis diversificados (es el caso de encuestas periódicas como los barómetros del CIS en España, la encuesta del Eurobarómetro, la encuesta europea de valores, etc.).

3.3. DE LOS CONCEPTOS A LAS VARIABLES

Con el término «concepto» nos referimos al *contenido semántico* (es decir, al significado) *de los signos lingüísticos y las imágenes mentales*. Su origen etimológico (del latín *cum capio* = tomar de manera conjunta), expresa la acción de ordenar varios elementos bajo un único acto de pensamiento, y el acto de extraer de la inmediatez de las impresiones sensibles y las representaciones concretas una abstracción con significado universal. Es el instrumento que permite al hombre conocer y pensar, y es el fundamento de toda disciplina científica, que consiste en conocer mediante conceptos universales.

De esta definición se desprende que «el término tiene un significado muy general, y puede incluir todo tipo de signos o procedimientos semánticos, sea cual sea el objeto al que se refiere, abstracto o concreto, cercano o lejano, universal o particular. Se puede tener un concepto de una mesa o del número tres, del hombre o de Dios, del género y de la especie [...] de un periodo o una institución históricos (el Renacimiento o el Feudalismo)» [Abbagnano, 1971, 146]. Los conceptos, además, pueden hacer referencia a construcciones mentales abstractas que no se pueden observar de manera directa, como el poder, la felicidad o la clase social, o bien referirse a entidades concretas y observables, como la flor o el obrero.

En las páginas anteriores hemos definido la *hipótesis* como una *conexión entre conceptos* o una *proposición que implica una relación entre dos o más conceptos*. La hipótesis de que la tasa de suicidios es más elevada entre las personas más instruidas implica un nexo, una relación, entre el concepto de suicidio y los conceptos de formación o religión. La hipótesis de que los valores postmaterialistas están más difundidos entre los jóvenes de las clases burguesas que entre los de la clase obrera establece una relación entre el concepto de valor y el de clase social.

Podemos decir, por tanto, que los conceptos son los «ladrillos de la teoría», y que mediante la operacionalización de los conceptos realizamos la traducción empírica de una teoría. El concepto tiende un puente que conecta la orilla teórica con la orilla del mundo empírico observable. Como escribió Blumer, «la teoría tiene sentido para la ciencia sólo en la medida en que logra conectarse con el mundo empírico de manera provechosa; los conceptos son los únicos instrumentos que permiten establecer tal conexión» [Blumer, 1969, 143].

Si la teoría es una red de conexiones entre entidades abstractas representadas por los conceptos, cuando estas entidades se hagan concretas, toda la red teórica se concretará también, y será posible establecer las mismas conexiones entre los conceptos transformados en entidades observables. Si la hipótesis teórica es que los valores postmaterialistas están más difundidos en las sociedades ricas, cuando estemos en condiciones de medir empíricamente la presencia de los valores postmaterialistas y la riqueza en distintas sociedades, podremos también verificar la validez de la teoría, observando si los datos de los dos conceptos operacionalizados muestran una correlación positiva.

¿Cómo se operacionaliza un concepto? Pensemos en los conceptos típicos de las ciencias sociales: poder, clase social, autoritarismo, participación electoral, desviación, subdesarrollo, etc. ¿Cómo podemos transformar estos conceptos en entidades observables?

El primer paso del proceso de traducción empírica de un concepto consiste en aplicarlo a objetos concretos, es decir, *convertirlo en atributo o propiedad de objetos concretos*, que llamaremos *unidades de análisis*. El concepto de poder, por ejemplo, puede ser una propiedad de unidades de análisis como cargos empresariales, cargos políticos (consejero municipal, funcionario de un partido, diputado, etc.). Así mismo, los conceptos de clase social y autoritarismo pueden ser propiedades de individuos; los conceptos de participación electoral y desviación pueden caracterizar a individuos o unidades territoriales (en este caso en forma de porcentajes de votantes y tasas de homicidios); el concepto de subdesarrollo puede ser una propiedad de los países. Nótese que estos conceptos del análisis social no tienen por qué ser necesariamente complejos: el género y la edad, el lugar de residencia y el tiempo empleado para ir de casa al trabajo también son ejemplos de conceptos más fáciles de medir.

Estas propiedades se presentan en distintos grados en los diferentes objetos, es decir, varían en las distintas unidades de análisis. Por ejemplo, unos cargos empresariales tienen más poder que otros, la clase social varía entre los distintos sujetos, la participación electoral y la criminalidad varían en los diferentes municipios.

El segundo paso para operacionalizar el concepto-propiedad consiste en formular una *definición operativa* de mismo, es decir, *establecer las reglas para su traducción en operaciones empíricas*. Por ejemplo, se puede definir operativamente el poder de un cargo empresarial por el número de personas que tiene a su cargo; el autoritarismo se puede operacionalizar con las respuestas a preguntas sobre actitudes, con puntuaciones entre 0 y 5; la participación electoral y la criminalidad de un municipio se pueden operacionalizar con el porcentaje de votantes sobre el número de electores en unas elecciones concretas y el tanto por mil de delitos sobre el total de la población.

El tercer paso consiste en *aplicar las mencionadas reglas a los casos concretos estudiados*: ésta es la fase de operacionalización en su sentido más estricto. La definición operativa se hace en el plano teórico, con la operacionalización se pasa a la práctica. Como escribió Marradi, la definición operativa es un «texto», mientras que la operacionalización es un «hecho».

La propiedad operacionalizada se denomina *variable*, y las *modalidades* son los «estados» operacionalizados de la propiedad. A cada modalidad se le asigna un *valor* simbólico distinto, que suele ser un número. Por ejemplo, el concepto de «nivel cultural» se puede representar con la propiedad «nivel de estudios» referida a los individuos, que adoptará diferentes estados en los distintos sujetos estudiados; estos estados se clasifican en alguna de las cinco modalidades siguientes: «sin estudios», «graduado escolar», «enseñanza secundaria», «diplomatura» y «licenciatura»; a cada una de estas modalidades les asignamos los *valores* 0, 1, 2, 3, 4.

Llegados a este punto conviene hacer una aclaración sobre el término «operacionalizar». En el lenguaje corriente habríamos dicho «medir», puesto que la medición es el proceso de asignación de valores numéricos a los estados de una propiedad. De hecho, la definición de medición que se suele usar en todos los manuales de metodología procede de las formulaciones de Stevens en los años cuarenta, según las cuales «la medición consiste en la atribución de números a objetos o acontecimientos siguiendo determinadas reglas» [Stevens, 1946].

En realidad, como ha señalado Marradi [1981, 602 sq.], cuando no se dispone de unidades de medida es inadecuado hablar de medición: en estos casos el paso de la propiedad a la variable es una operación distinta a la *medición* (por ejemplo, medimos la edad con la unidad de medida del año solar), que, según los casos, es una *clasificación* (por ejemplo, clasificamos la nacionalidad según la nación), una *ordenación* (por ejemplo, ordenamos las profesiones en función de su prestigio social), o un *cómputo* (por ejemplo, contamos el número de hijos de una persona).

Sin embargo, en el lenguaje común no existe un término único para designar esa operación general de medición-ordenación-cómputo-clasificación. Las ciencias naturales, donde las unidades de medida casi siempre son definibles, han impuesto el término «medir» para referirse a la operación general. Pero para la clasificación o la ordenación el término «medir» resulta inadecuado, por ello hemos empleado el término «operacionalizar». Este término tiene un sentido amplio y se emplea con el significado general de «traducir del lenguaje teórico al lenguaje empírico», pero en sentido estricto se refiere al *paso de las propiedades a las variables*, que, en el proceso más amplio que conecta la teoría y la investigación, representa el puente

crucial entre las dos vertientes, y que puede esquematizarse del siguiente modo:



3.4. UNIDAD DE ANÁLISIS

La «unidad de análisis», como ya se ha indicado, representa el objeto social al que se refieren las propiedades estudiadas en la investigación empírica. Una reflexión teórica no necesita definir unidades de análisis exactas. Pensemos, por ejemplo, en la interpretación de la rebelión social y el radicalismo político en términos de «privación relativa». Según esta interpretación, «la insatisfacción con el sistema de privilegios y recompensas de una determinada sociedad no [es proporcional] al grado de desigualdad que afecta a sus miembros» [Runciman, 1966; 3], sino que depende más bien del sentido de privación que tengan dichos miembros, que a su vez deriva de los «grupos de referencia» y comunidades con los que los individuos se comparan, y las expectativas que tienen. Por ello, «la manifestación de descontento es poco frecuente en periodos de privaciones constantes, mientras que tiende a aumentar [...] cuando se entrevé una posibilidad de mejora [...] [y] es probable que las revoluciones se produzcan en periodos de bienestar creciente» [*ibidem*, 21]. Podría desarrollarse una teoría similar sobre la rebelión en el plano de la reflexión filosófica, histórica, social y psicológica, aplicándola a distintos contextos, problemáticas y hechos históricos. Pero en el momento en que queramos comprobar esta teoría mediante una investigación cuantitativa concreta, al definir el «diseño de la investigación», debemos definir ante todo la unidad de análisis. Ésta podría estar representada por el *episodio* de rebelión social, como en el estudio de Gurr y sus colaboradores [Gurr, 1972, 92-98], que recopilaron de manera sistemática datos sobre actos de violencia política y protesta social en el periodo 1961-65 en 38 países, en concreto, una serie de informaciones publicadas en la prensa. Por ejemplo, el número de participantes y víctimas, la duración de la agitación, el tipo de personas implicadas, los motivos del descontento, las modalidades de la protesta, las características del contexto social (tipo de régimen político, variables económicas, libertades civiles, etc.), llegando a registrar más de 1000 episodios. Todo ello con el objetivo de comprobar si los casos de conflicto y violencia estaban ligados a situaciones de cambio social que habían originado expectativas no satisfechas, y en general a situaciones de privación relativa más que de privación absoluta.

La unidad de análisis de una investigación orientada a comprobar la teoría mencionada también podría estar representada por un *núcleo territorial*. Aceptando una sugerencia de Tocqueville, para quien la revolución francesa se desarrolló en las zonas más ricas del país, puesto que «las partes de Francia que habían de ser el foco principal de esa revolución son precisamente aquellas donde el progreso es mayor» [Tocqueville, 1856], se podría realizar una investigación tomando como

unidad de análisis la región. Se trataría de encontrar un indicador de bienestar económico y fuerza del impulso revolucionario para las regiones de Francia en la época de la revolución, y relacionar ambas variables.

También se podría tomar como unidad de análisis al *individuo*. En 1962, Runciman entrevistó a unos 1400 ciudadanos de Inglaterra y Gales, utilizando preguntas abiertas y cerradas, para identificar los grupos de referencia de cada sujeto, la clase a la que consideraba que pertenecía, el grado de satisfacción con su posición social, las aspiraciones insatisfechas, es decir, las componentes de la privación relativa, junto con las orientaciones políticas, para ver si existía una correlación entre las dos variables.

Hemos dicho que la traducción empírica de un concepto (que es abstracto por definición) se hace mediante la asignación de este concepto a un objeto social concreto («unidad de análisis»), como si fuera una propiedad de dicho objeto. En una primera aproximación, podemos identificar los siguientes tipos de unidades de análisis: el *individuo*, el *conjunto de individuos*, el *grupo-organización-institución*, el *acontecimiento* y el *producto cultural*.

3.4.1. Los distintos tipos de unidad de análisis

La unidad de análisis más frecuente en las investigaciones sociales está representada por el individuo. Por otra parte, como escribe Galtung, «la sociología se define a menudo como la ciencia de la interacción social, de lo que se deriva que la unidad de los análisis sociológicos debería ser el actor social [...], y en la mayoría de los análisis sociológicos la elección recae sobre el actor social por excelencia, el individuo humano» [Galtung, 1967, 37].

También es bastante frecuente que la unidad de análisis esté representada por un *colectivo*. Este colectivo puede ser un conjunto de individuos o un grupo-organización-institución. El ejemplo más común de *conjunto de individuos* son las fuentes estadísticas oficiales, basadas en agrupaciones de individuos por territorios (municipios, provincias, comarcas, regiones...). En este caso las variables proceden de operaciones matemáticas realizadas sobre variables registradas de forma individual (por ejemplo, renta media, porcentaje de desempleados, etc.). Si hacemos una distinción, como a veces se hace, entre *unidad de análisis* y *unidad de registro*, en este caso la unidad de registro se sitúa en un nivel inferior respecto a la unidad de análisis, y está representada por el individuo.

En un *grupo-organización-institución* la mayoría de las variables está representada por lo que llamaremos «variables colectivas estructurales» o «globales», y la unidad de registro es el propio colectivo. Pensemos en un estudio sobre centros escolares, en el que las variables se refieren al tipo de gestión (privada/pública), la realización (o no) de programas experimentales, el tipo de perfil social al que se dirige la escuela, el número de clases, el número de alumnos, la presencia de mujeres en el profesorado, el índice de aprobados, etc. Aunque algunas variables (como las dos últimas citadas) hacen referencia a ámbitos individuales subyacentes, los datos se recogen de forma colectiva (por ejemplo, mediante una entrevista con el director y una consulta de los archivos del instituto), por lo que la unidad de análisis

y la unidad de registro coinciden.

En investigación social estas unidades de análisis son bastante frecuentes. Pensemos en *grupos* como familias, asociaciones, sectas religiosas, grupos étnicos, bandas juveniles, pandillas, y *organizaciones-instituciones* como los sindicatos, partidos, organizaciones laborales, unidades administrativas locales, instituciones públicas, hasta sociedades enteras y los propios países.

Una cuarta unidad de análisis es el *acontecimiento*. Ya hemos citado la investigación de Ted Gurr sobre los episodios de rebelión política. Otro acontecimiento podría ser la elección política: podemos citar como ejemplo la investigación de Bartolini y Mair [1990], que estudiaron las elecciones de las democracias europeas celebradas entre 1885 y 1995, en relación con la inestabilidad electoral, el voto de clase, la participación electoral, etc.; cada una de estas variables, obviamente, representa una propiedad de la unidad de análisis «acontecimiento electoral». Otros acontecimientos podrían ser huelgas, guerras, golpes de estado, procesos judiciales, ceremonias religiosas, campañas electorales, manifestaciones de protesta, etc.

El último tipo de unidad de análisis es lo que podríamos llamar *representación simbólica-producto cultural*. Pensemos en toda la corriente del análisis de contenidos, donde la unidad de análisis está representada casi siempre por mensajes de los medios de comunicación de masas, escritos, orales o audiovisuales: artículos de periódicos, textos literarios, programas electorales, discursos políticos, fotografías, programas de televisión, espectáculos de teatro, reportajes.

Finalmente queremos precisar que llamamos *observaciones* a los ejemplares de una determinada unidad de análisis incluidos en una investigación. La «unidad de análisis» es una definición abstracta, que designa el tipo de objeto social al que se refieren las propiedades (como el elector o el episodio de protesta). Esta unidad se localiza en el tiempo y en el espacio, definiendo «la población de referencia» de la investigación (por ejemplo, los electores españoles en las elecciones generales de 2000; los episodios de protesta política registrados en España entre 2002 y 2003). Esta población de referencia puede estudiarse de manera integral, por ejemplo, la investigación sobre los gobiernos que se han sucedido en España desde la llegada de la democracia en 1977 hasta el año 2004, o de manera parcial. A menudo se intentará extraer una muestra casual, otras veces se podrán usar distintos criterios de selección. Las observaciones son los ejemplares de esa unidad de análisis que se estudian, sobre los cuales se recopilan datos. Mientras que la unidad de análisis es singular y abstracta, las observaciones son múltiples y concretas, y constituyen el objeto específico de la investigación empírica. En definitiva, en una matriz de datos las observaciones son todos y cada uno de los números incluidos en las casillas.

3.5. VARIABLES

Como ya hemos dicho, una *variable* es un *concepto operacionalizado*. Más en concreto, consiste en la *propiedad operacionalizada* de un objeto, ya que para poder operacionalizar el concepto, debemos aplicarlo a un objeto, y de este modo se convierte en propiedad. Entre concepto, propiedad y variable existe la misma diferen-