

Indicators: Syntactic Vision and Semantic Vision: First Part (Indicadores: la Visión Sintáctica y la Visión Semántica: Primera Parte)

Alberto Marradi¹

²doi:10.23756/sp.v5i2.396



Abstract

In this essay several aspects of the concept 'indicator' are criticized, as well as the technical consequences of those misconceptions. The most relevant among them is the pretence of measuring the validity of an indicator, e. g. the degree of semantic correspondence between the indicator and the concept which it is supposed to refer to.

Keywords: indicator, operational definition, semantics, validity, measurement.

Palabras claves: indicador, definición operacional, semántica, validez, medición

Premisa

En este ensayo me propongo criticar una serie de tesis explicitadas y de supuestos acerca de la naturaleza y la función de los indicadores y, por consiguiente, una serie de procedimientos que se derivan de esas tesis y supuestos, en particular:

- 1) el supuesto según el cual se hace uso de ellos solo en las ciencias sociales;
- 2) la burda tendencia a confundirlos con las definiciones operativas; o
- 3) la tesis casi universal según la cual todas las propiedades de gran generalidad e importancia teórica necesitan indicadores;
- 4) el abanico de interpretaciones que se dieron acerca de la relación entre indicador y concepto indicado, todas inspiradas por el deseo de conjurar el

¹ Emeritus, Universidad de Florencia; alkmar@libero.it.

² ©Alberto Marradi. Received: 20-09-2017. Accepted: 27-12-2017. Published: 31-12-2017.

odiado término ‘semántica’, que remite a la disciplina justamente considerada como la más peligrosa enemiga de la visión objetivista de la ciencia;

5) la interpretación predominantemente o exclusivamente sintáctica de la validez, es decir, del grado de correspondencia (me resulta difícil no usar el adjetivo ‘semántica’) entre indicador y concepto indicado;

6) la consiguiente pretensión de medir la validez con varios coeficientes;

7) la idea según la cual establecer una relación de indicación es ilegítimo, porque cada concepto de propiedad tiene su autonomía;

8) la idea según la cual la relación de indicación tiene que ser unívoca, y que el hecho de que investigadores diferentes usen indicadores diferentes para el mismo concepto es una demostración de la inmadurez de las ciencias sociales.

A cada punto a criticar dedicaré un párrafo. Luego agregaré unas propuestas de elección de indicadores por algunos conceptos fundamentales de las ciencias sociales.

1 Indicadores en la vida cotidiana*

En la vida cotidiana se usan continuamente indicadores: lo saben muy bien los estafadores que se presentan ante viejas viudas para proponerles negocios super-ventajosos: si no se visten como caballeros de industria, la señora desconfiará de ellos y, en vez de confiarles sus ahorros, se despedirá amablemente y pronto llamará a su amiga o a su abogado. También lo saben los ladrones que planean desvalijar un banco: solo si entran perfectamente vestidos de saco y corbata, los clientes y los taquilleros los considerarán parte del ambiente y se darán cuenta de ellos únicamente cuando estén siendo apuntados por una pistola.

Naturalmente, no siempre los indicadores son utilizados en tanto engañosos como en estos casos. Tomemos el ejemplo de una joven cuyo rostro enrojece: el hecho puede ser considerado, por diferentes observadores, como indicador de pudor frente a un cumplido, de naturaleza retraída, de empacho frente a una pregunta delicada, de toma de conciencia de estar diciendo una mentira, de sensación de culpa, de esfuerzo intelectual frente a un problema difícil e incluso de esfuerzo físico. El contexto puede excluir con seguridad la última interpretación y también otras, pero casi nunca puede eliminar todas las alternativas.

Aquí se encuentra una característica que pertenece también a los indicadores usados en las ciencias sociales y que llamaremos «pluralidad desde

* Para los párrafos 2 y 3 he utilizado las notas tomadas por los licenciados Leandro Eryszewicz y Andrea Pereyra Barreyro, que asistieron a mi curso de metodología en la Universidad de Buenos Aires en el periodo marzo-abril del 2013. Les agradezco por su valioso trabajo y por la revisión del texto entero de este ensayo.

arriba».³ Una propiedad puede ser considerada indicador de dos o más o propiedades, por diferentes observadores o incluso por el mismo. Por ejemplo, la tasa de divorcios puede ser considerada un indicador de secularización o de modificación de las estructuras familiares, etc.

Puede también pasar que el rubor interese de por sí. Por ejemplo: si una estilista se da cuenta de que una joven clienta tiene una tendencia a sonrojarse, le puede proponer algún traje de un color que sea acorde con un rojo pálido. Aquí también hay un punto de contacto con los indicadores usados en las ciencias sociales: para un investigador que estudia la conducta de la policía el número de robos en un distrito tiene un interés en sí; a otro, le interesará en tanto indicador del nivel de inseguridad urbana; a un tercero, como indicador de la difusión de la anomia o del incumplimiento de las leyes. La asertividad de un joven puede interesar por sí misma o como indicador de un complejo de inferioridad; la tasa de divorcios puede interesar en sí o como indicador de secularización; el porcentaje de votos a un partido de oposición puede interesar por sí misma o como indicador de insatisfacción con la política del gobierno, etc.

1.1. Como pasa con muchos conceptos y meta-conceptos,⁴ las ciencias sociales tomaron el concepto de indicador de la vida cotidiana y lo formularon con mucho retraso, debido naturalmente a su tardía consolidación. Se puede afirmar que el concepto (pero no el término) ya les era familiar a los sociólogos y estadísticos positivistas franceses de 1800: a Villermé (1840) tanto como a Quetelet (1869). A finales del siglo, Durkheim escribió de «indices extérieurs que symbolisent des faits intérieurs et des phénomènes moraux» (1893, 1896). En el nuevo siglo, el italiano Nicéforo usó el término ‘índice’ en una acepción parecida a la que actualmente se atribuye a la expresión ‘indicador social’ (1921).

El operacionista norteamericano Stuart Carter Dodd (1942) es el primer sociólogo que usó corrientemente el término en su sentido técnico actual. Finalmente, Paul Felix Lazarsfeld (1958) codificó el uso actualmente predominante en sociología y estableció la distinción entre ‘indicador’ e ‘índice’.

1.2. Como se vio, tanto en el lenguaje ordinario como en el lenguaje técnico de las ciencias sociales el término designa algo que puede ser interesante por sí mismo (el rubor/la tasa de robos) o como fuente de informaciones sobre varias otras cosas (el pudor, el esfuerzo/la inseguridad, la anomia).

³ Véase §8, en la segunda parte de este ensayo.

⁴ Uso el término ‘meta-conceptos’ para conceptos metodológicos, como indicador, medición, ley, etc.

Las diferencias entre el uso ordinario y el uso técnico dependen de la naturaleza de la matriz de los datos, la cual no se usa en la vida cotidiana, mientras que es el instrumento básico en la aproximación estándar a la investigación social. En su paso de ser un término de ordinario a ser uno técnico sufrió los cambios de significado necesarios para ser adaptado a la forma estándar de la investigación: relevar los estados en una serie de propiedades de una multitud de objetos del mismo género (seres humanos, provincias, estados, etc.) en un ámbito espaciotemporal dado y transformarlos de manera que puedan ser registrados en el vector de una matriz de datos.

Esa transformación se realiza con una serie de operaciones, siguiendo convenciones y reglas: el complejo de estas operaciones, convenciones y reglas ha tomado el nombre de «definición operativa».

2 Indicadores y definiciones operativas

Hay una cierta tendencia, especialmente en autores norteamericanos, a desconocer las diferencias entre indicadores y definiciones operativas, trastocando todo el campo semántico. Baste con dos ejemplos:

«Los conceptos son más abstractos que las definiciones operativas, y más que una definición operativa pueden ser indicadores del mismo concepto» (Reynolds, 1971, p. 143).

«Hay que convertir a los conceptos en indicadores operativos, así llamados porque hacen explícitas las operaciones con las que los fenómenos son expresados en datos científicamente útiles» (Singer, 1982, p. 194).

La definición operativa consiste en un conjunto de reglas, convenciones y prácticas que permiten convertir una propiedad de objetos (que pueden ser seres humanos, agregados territoriales, etc.) en una variable al interior de la matriz de datos. En detalle, se trata de todas las operaciones (tanto intelectuales como materiales) que permiten pasar de un estado de cualquier objeto cognitivo en una propiedad a un signo (generalmente un número), ello en un vector-columna que represente aquella propiedad en una matriz de datos. Se entiende fácilmente que se trata de un instrumento básico en toda investigación de tipo estándar: sin definiciones operativas no habría matrices de datos, y sin matrices de datos no habría ni análisis descriptivo (con todas esas sofisticadas técnicas que hacen las delicias de los «cuantitativos»), ni tampoco estadística inferencial, con sus pretendidos «universos» (es decir, conjuntos con infinitos miembros, imposibles en el mundo real), los cuales son artilugios terminológicos para ocultar las peticiones de principio (*affirming the consequent*) de sus veneradas fórmulas.

Como sabe quienquiera que haya realizado una investigación directamente — y no a través de consultoras o colaboradores — a veces se encuentran dificultades prácticas al ejecutar una definición operativa, así que el investigador o el metodólogo tienen que valerse de su arte⁵ para soslayarlas. Pero otras veces se encuentran dificultades conceptuales que no son eludibles. Un ejemplo nos permitirá comprenderlo cabalmente: imaginemos que la propiedad de interés sea la libertad política, la unidad de análisis sea el Estado y el ámbito América. Una definición operativa no podría ser muy diferente de la que sigue:

- a) Buscar al señor Brasil, a la señora Argentina, el señor Canadá, etc., y preguntarles cuál es el nivel de libertad política de su ciudadanía.
- b) Solicitarles que respondan la anterior pregunta eligiendo uno de los siguientes niveles: ninguno; escaso; insuficiente; suficiente; bueno; óptimo.

Como puede observarse, esto resulta casi ridículo, porque el señor Brasil, la señora Argentina, el señor Canadá, etc., no existen. Y no es una plausible solución alternativa poner la misma pregunta a los jefes políticos de cada Estado porque su respuesta sería poco confiable: se registraría la imagen que los presidentes, o dictadores, quieren dar del nivel de libertad política de sus países, que es una cosa distinta del nivel en sí mismo.

Como consecuencia general, no se puede imaginar una definición operativa directa cuando:

- a) la unidad de análisis es algo que no se puede interrogar (una institución, un agregado territorial, etc.);
- b) los estados en la propiedad investigada no se puedan registrar directamente (como, por ejemplo, la superficie) u obtener de documentos oficiales (como se puede hacer para el número de habitantes, el tipo de sistema electoral, el número de votos de un partido, etc.).

Por otro lado, cuando el objeto cognitivo es un individuo que se puede interrogar o una definición operativa directa (que pone bajo forma de pregunta el concepto de propiedad que interesa investigar) siempre es posible; aunque esto no implica que sea siempre aceptable. Para comprobarlo, imaginemos ahora que nuestra unidad de análisis sea el varón adulto y la propiedad que interesa sea su grado de autoritarismo. Una definición operativa directa incluiría preguntas como: ¿cuál es a su parecer su grado de autoritarismo?, y respuestas del tipo: inexistente / bajo / medio / alto. Pero a veces los caminos sencillos no conducen al lugar debido; en este caso, no se registraría el grado de autoritarismo, sino la propensión de cada varón a declararse autoritario — una propensión rara entre los varones en general y quizás rarísima entre los efectivamente autoritarios. No se descarta la posibilidad de que la variable así construida tuviese una relación inversa con el efectivo autoritarismo, ya que

⁵ Véase Marradi (2002).

unos sujetos débiles y sumisos podrían declararse “bastante autoritarios” como para presentar una imagen más gratificante de sí mismos.

No es un ejemplo aislado como puede parecer. Imagínese que preguntamos a los miembros de una muestra, esta vez de varones y mujeres, cuán sinceros, leales o generosos son. No se descarta la posibilidad de que algunos admitan no gozar de una u otra calidad o, incluso, de ninguna de las tres. Pero parece plausible que la distribución de las respuestas sería sesgada en dirección positiva, por la tendencia de muchos a auto-valorarse positivamente, de buena o de mala fe. Imagínese que preguntamos a los mismos miembros si son particularistas, esquizoides o maniaco-depresivos: en este caso la distribución de las respuestas sería fortuita, porque muchos hombres y mujeres de la calle no conocerían el significado de esos términos, pero al no admitirlo contestarían al azar. Y finalmente imagínese que les preguntamos si evaden el fisco, consumen drogas, son bisexuales, etc. También en estos casos la distribución de las respuestas sería sesgada, pero en dirección de lo que se considera social o moralmente aceptable.

En consecuencia general: cuando la unidad de análisis es un ser humano, se puede imaginar una definición operativa directa; pero esa no será confiable (en el sentido de que producirá muchos datos no fidedignos) si la propiedad investigada:

- a) tiene respuestas socialmente deseables⁶ o indeseables;
- b) es algo familiar al científico pero no al lego;
- c) es algo que la moral dominante en la comunidad del entrevistado considera reprochable.

De tal forma, en todos los casos listados arriba una definición operativa no es posible o no parece confiable. Esto no conlleva una renuncia a recoger informaciones sobre esa propiedad. Para hacerlo, es necesario buscar una (o, como veremos, preferentemente más) propiedades que:

- a) admitan aceptables definiciones operativas directas;
- b) tengan una fuerte relación semántica con la propiedad que interesa.

Allí empieza — y acaba — la función de los indicadores en la investigación social.

3 ¿Todas las propiedades de gran generalidad necesitan indicadores?

Muchos autores de gran importancia en las ciencias sociales⁷ escriben que los indicadores son necesarios cuando el concepto es general o tiene gran

⁶ Es decir, respuestas que un entrevistado tiende a elegir para dar una imagen de sí (Goffman, 1956) más agradable frente al entrevistador, a quien considera un “otro generalizado” (Blumer, 1948).

importancia teórica. Sin embargo, un estadístico dotado de un singular sentido común observó que «algunas variables teóricamente importantes pueden ser fácilmente operativizadas: la edad, el sexo, la raza, la confesión religiosa» (Blalock, 1961, p. 163). A esta corta lista se pueden agregar la nacionalidad, el lugar de residencia, el título de estudio, el estado civil, la condición laboral, el tipo de ocupación y, para el caso de las unidades territoriales, la superficie, la población, la forma de régimen político, etc. (tabla 1).

Tabla 1. *Propiedades de gran generalidad y relieve teórico que no necesitan indicadores*

		Unidad individual	Unidad territorial
Ejemplos de Blalock		Edad, sexo, raza, creencia religiosa	—
Otros ejemplos		Nacionalidad, provincia de nacimiento/residencia, dimensiones del lugar de nacimiento/residencia, título de estudio, estado civil, condición ocupacional, tipo de ocupación, etc.	Superficie, población; cuotas de la población en las diferentes franjas de edad/condiciones ocupacionales/sectores de ocupación; tipo de régimen, número de votos para un partido

Cada propiedad de la tabla 1 tiene un alto nivel de generalidad y puede tener gran relevancia teórica; sin embargo, de ellas se pueden dar definiciones operativas y, por tanto, no necesitan indicadores.

Puede afirmarse, entonces, que los autores mencionados en nota 4 cometieron un clásico error lógico: de la observación indiscutible que los indicadores son conceptos más específicos de los conceptos que indican,⁸ sacaron la conclusión de que todos los conceptos generales necesitan indicadores. Pero esto es claramente un *non sequitur*, es decir, una deducción equivocada porque la premisa no conlleva la conclusión.

La conclusión que inevitablemente se desprende de cualquier análisis no superficial de lo que pasa efectivamente en la investigación social es que se recurre a indicadores si, y solo si, no es posible dar una satisfactoria definición operativa de una propiedad que interesa, independientemente de su nivel de generalidad, relevancia teórica u otra cualidad. Si no puede relevar directamente los estados en la propiedad \mathbb{A} , el investigador piensa en una o más propiedades

⁷ Los sociólogos Merton en 1948 y Galtung en 1967; los politólogos Przeworski y Teune en 1970; el epistemólogo Nowak en 1976; los metodólogos Lazarsfeld y Barton en 1951; McKennell en 1973; Carmines y Zeller en 1979, así como muchos otros más.

⁸ El motivo de esto se verá en el §8, en la segunda parte de este ensayo.

cuyos estados se puedan relevar con una definición operativa y que él juzga semánticamente asociada(s) a la propiedad A.

Por tanto, el indicador no es una forma particular de definición operativa, ni algo intercambiable con ella, ni algo que la sustituye (como aparece en algunos textos de metodología). Es nada más y nada menos que un concepto de propiedad el cual, como tal, solo puede sustituir a otro concepto de propiedad; y lo sustituye por decisión que un investigador o grupo de investigación toman al darse cuenta que el segundo no admite una confiable definición operativa que permita convertirlo en una variable. Como se dijo anteriormente (§1.2), la función de la definición operativa en las ciencias sociales es relevar los estados en una serie de propiedades de una multitud de casos y transformarlos de manera que puedan ser registrados en el vector de una matriz de datos. Ya que uno de los supuestos del método de la asociación, que caracteriza la aproximación estándar en las ciencias sociales (Marradi 2007, §4), es que las variables que ocupan los vectores-columnas deben ser referibles a la unidad de análisis elegida, y que todos los casos —ejemplares de esta unidad de análisis— deben en principio tener estados en todas las propiedades que son transformadas en variables por la definición operativa, un indicador tiene que cumplir con ese supuesto: si una porción bastante amplia no puede en principio⁹ tener estados sobre la propiedad, esta no puede ser elegida como indicador.

Un ejemplo mostrará mejor lo que se quiere decir. Si al investigar una muestra aleatoria de ciudadanos se elige, entre los indicadores de estatus social, el tipo de estación de esquí preferido por el individuo, eso no se puede aceptar como indicador de estatus social porque en su mayoría los ciudadanos, o no saben esquiar, o no tienen medios para frecuentar regularmente estaciones de esquí.

Un indicador que se descarte por este motivo se denomina «no vectorial». Naturalmente, poder producir un vector completo de datos es una condición necesaria pero no suficiente: la condición esencial es que un investigador o equipo de investigación¹⁰ perciban una relación semántica suficientemente estricta entre el indicador y el concepto indicado.

4 La naturaleza semántica de la relación de indicación

En los manuales de metodología se encuentran muchas posturas diferentes sobre la naturaleza de la relación entre indicadores y conceptos indicadores –

⁹ Se subraya la expresión «en principio», considerando que se pueden presentar «casos faltantes» dependiendo de situaciones accidentales.

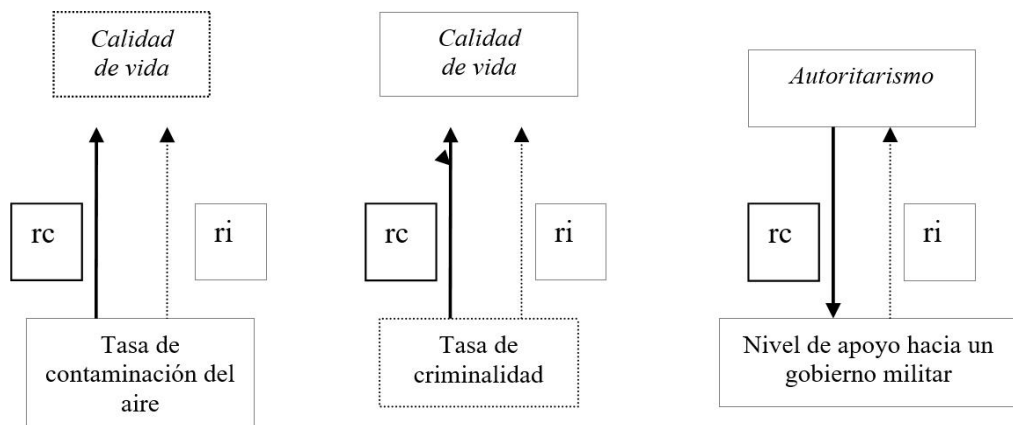
¹⁰ Naturalmente, esta percepción es sometida al juicio de la comunidad científica. Si la elección no es obvia –como en efecto lo es en la mayoría de los casos–, sería de esperar que el investigador la comente para justificarla.

relación que raramente se denomina—. Un nombre/rótulo adecuado me parece «relación de indicación». A continuación, se examinarán dichas posturas una tras otra.

4.1. A partir de Stevens (1951) y siguiendo con Blalock (1961), Sullivan (1974), Smelser (1976) y muchos otros, se acostumbra decir que la relación de indicación es una relación causal, generalmente sin establecer si el concepto indicado es la causa o el efecto del concepto indicador. Como veremos enseguida, hay casos en que los conceptos que fueron y son elegidos como indicadores se pueden considerar razonablemente causas del concepto que estarían indicando, así como existirán casos en que se puedan considerar como efectos; y, por otro lado, en algunos casos los indicadores que fueron y son elegidos no se pueden considerar ni causa ni efecto del concepto indicado. Aun en los casos en los que se puede imaginar que esta relación exista, establecer su dirección no es siempre tan obvio como parece.

Algunos ejemplos: la tasa de contaminación del aire o la de criminalidad son algunas causas de la percepción de la calidad de vida en un municipio, y a menudo se eligen como indicadores de dicho concepto. Al contrario, la tasa de apoyo hacia un gobierno militar puede considerarse un efecto de una actitud autoritaria más general del sujeto, y frecuentemente se elige como indicador de autoritarismo (figura 1).

Figura 1. *La relación de indicación puede ir de la causa al efecto o del efecto a la causa*



Nota. **rc** indica una relación causal, mientras que **ri** una relación de indicación.

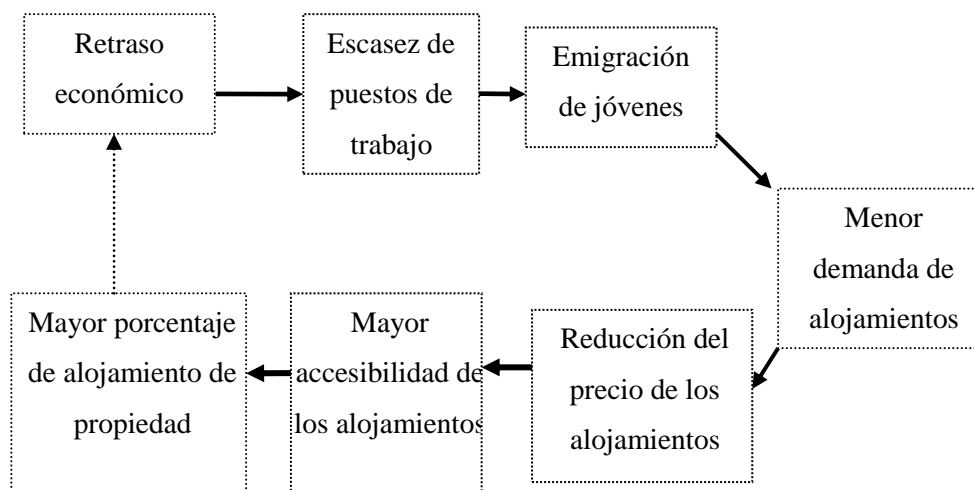
La relación causal puede ser indirecta, pasando por una cadena con muchos anillos. Un caso llamativo es el de la proporción de casas poseídas por la familia que las habita. A primera vista, esta propiedad parece un posible indicador de bienestar económico: cuanto más rica es una familia, tanto más debería estar en condiciones de comprar la casa donde habita. Pero en los diagramas (*plots*) producidos por varios análisis factoriales desarrolladas sobre datos socioeconómicos-demográficos relativos a comunas de varias regiones italianas (Marradi 1978, 1979, 1981), esta variable puntualmente aparecía ubicada en el mismo racimo de otras variables vinculadas al subdesarrollo (porcentaje de analfabetos, porcentaje de ocupados en el sector primario, etc.). De manera que en ese momento controlé si ello dependía de un artefacto del análisis factorial – como a veces sucede – pero en cada situación resultaba una alta correlación negativa entre el ingreso per cápita del municipio y la proporción de casas poseídas por la familia que las habitaba. Me puse a reflexionar en la cuestión y comprendí que esa relación no era aleatoria, sino el efecto de una verdadera cadena causal: 1) en los municipios de la montaña – o aislados de alguna otra forma – las oportunidades de trabajo son limitadas; 2) por ese motivo, los ingresos medios son bajos; 3) los jóvenes tienden a desplazarse a municipios urbanos y céntricos; 4) esta migración interna reduce la demanda de alojamiento en los municipios aislados y pobres y la acrecienta en los municipios ricos y céntricos; 5) por eso, los precios de los alojamientos tienden a bajar en los primeros y a subir en los segundos; 6) por consiguiente, incluso muchos habitantes pobres de municipios aislados tienen bastante dinero como para comprar la casa donde viven, mientras los habitantes más ricos de los municipios céntricos no consiguen hacer lo mismo.

Esta experiencia me confirmó que las ecuaciones *indicador = causa* o *indicador = efecto* pueden ser demasiado burdas y que el papel de la reflexión e interpretación de los hechos en campo es esencial. Al haber imaginado una cadena causal como la que aparece en la figura 2, en la cual cada eslabón resulta persuasivo, el investigador puede estar confiado en elegir el porcentaje de casas que son de propiedad de los que las habitan como indicador de subdesarrollo económico, mientras que si se postulaba cualquier relación causal directa los resultados del análisis, estos contradecían el sentido común.

Más sutilmente, se puede señalar que la tesis de que la relación de indicación es una relación causal confunde los planos. Supongamos que un investigador haya observado que muchos individuos autoritarios están a favor de un gobierno militar y piense que el autoritarismo cause tal preferencia. La relación causal que el investigador imagina entre el nivel de autoritarismo de los sujetos y sus respuestas a una pregunta sobre el gobierno militar *no es* la relación de indicación que él instituye entre los dos conceptos, sino que es la *justificación* de tal relación. Pero a veces no se logra directamente operativizar y se elige la primera como indicador de la segunda basándose en una simple

asociación estadística.

Figura 2. *Cómo una cadena causal muy articulada puede respaldar la elección de un indicador*



Nota. Las flechas sólidas indican relaciones empíricas; la flecha sutil indica una relación semántica.

A título de ejemplo se puede tomar un renombrado test psicológico: el *Minnesota Multiphasic Inventory* (Hathaway & McKinley, 1939). En ese test, muchas de sus 480 preguntas parecen bizarras; un ejemplo es: «¿Siente usted un prurito en sus dedos cuando abre un picaporte?» Es posible que los psicólogos hayan detectado relaciones estadísticas entre sensaciones tan específicas y extrañas e importantes propiedades psicológicas. Basándose en ellas, han decidido elegir una u otra sensación como un indicador de una u otra patología psíquica sin explicitar cuál sería el nexo causal entre las dos. La elección de tal conducta como indicador de tal patología está basada en una correlación observada y no problematizada.

Como decía, los planteamientos acerca de la naturaleza de la relación de indicación son variados, y muchos respetados metodólogos, psicólogos y sociólogos han propuesto sus versiones: Lazarsfeld (1958) y otros que lo siguen (por ejemplo Verba, 1969; Nowak, 1976) hablan de una relación probabilística. Pero esto es un burdo error gnoseológico, ya que una relación probabilística puede darse solo entre un evento (o clase de eventos) y otro evento (o clase de eventos).

También Galtung (1967), McKennell (1973) y Singer (1982) hablan de inferencia; esto también es un error gnoseológico, ya que la inferencia es un posible vínculo entre aserciones, nunca entre conceptos. Por su parte, Guttman (1950), Cronbach y Meehl (1955) y otros hablan de «muestreo de un universo de contenido». Pero este es otro error gnoseológico más, ya que el muestreo

supone que la población (y por tanto, la muestra) sea formada por individuos — u otros objetos físicos — delimitados, no por entidades vagas y no delimitadas como los conceptos.

Menos grave parece el error de los que hablan de una relación lógica o analítica (por ejemplo, Nowak, 1976; Sullivan & Feldman, 1979). Una relación lógica en efecto puede establecerse entre conceptos (por ejemplo, la relación entre un género y sus especies es de este tipo); pero la relación entre un concepto y sus indicadores no tiene la misma naturaleza que la relación género-especie, la cual es necesaria (todos los gatos son felinos, pero hay felinos que no son gatos). Cada investigador, o grupo de investigadores, la establece según su conocimiento de la población estudiada y de la literatura sobre el tema, sus experiencias en investigaciones previas y, en definitiva, según su preferencia. Otros investigadores (tanto como el mismo investigador en una ocasión posterior) pueden elegir otros indicadores para el mismo concepto, lo cual ocurre a menudo.

De la reseña de este párrafo se puede concluir que estos autores son proclives a suscribir a cualquier interpretación, incluso si es patentemente falaz, con tal de no pronunciar el término-tabú ‘semántica’ que — como decía en páginas anteriores — no es conciliable con una visión objetivista de la ciencia, evidentemente compartida, abiertamente o subrepticamente, por cada uno de ellos. El corazón de la visión semántica que se propone en este ensayo es la distinción más neta que sea posible entre una relación semántica entre conceptos y una relación empírica entre propiedades (o, si así se prefiere, una relación sintáctica entre vectores). Los metodólogos de orientación técnica tienden a confundir las dos cosas porque las herramientas que se usan para controlar empíricamente las dos relaciones son, en parte,¹¹ los mismos (coeficientes de correlación entre vectores). Pero técnicas formalmente iguales pueden ser empleadas para fines radicalmente diferentes.

Para esclarecer la diferencia entre un vínculo empírico y uno semántico se pueden considerar dos propiedades: la edad (que no necesita indicadores) y el conservadurismo (que sí los necesita, porque pocos entrevistados no-británicos admitirían ser conservadores). Supongamos que en una muestra se encuentra

¹¹ Pero no totalmente. Así, los psicólogos de la primera mitad del siglo XX usaban una forma de control de validez denominada «por grupos conocidos» (*known groups*), la cual era radicalmente distinta de una correlación entre vectores y se basaba en intuiciones psicológicas más que en cálculos estadísticos. Para ver si un test dado era un indicador válido, por ejemplo, de anomia, se lo aplicaba a una población conocida (habitualmente, los alumnos de una clase de psicología). El test era juzgado válido si el grupo de alumnos que el profesor juzgaba «anómicos» obtenía en el test un puntaje medio significativamente más alto que el resto de la clase (Thurstone & Chave, 1929; Frey, 1970). Técnicamente este procedimiento se asemeja a un análisis de la varianza, no a un análisis de correlación-regresión. Pero nunca se aplicaban los coeficientes asociados a este tipo de análisis (eta cuadrado, test F, etc.): el juicio no era confiado a una cifra singular sino remitido a la evaluación de psicólogos al tanto de la situación específica.

una correlación del 0.7 (muy alta para un sondeo) entre la edad y cualquier índice de conservadurismo; ¿de esto se derivaría que podemos considerar a la edad como un indicador de conservadurismo? Creo que incluso los partidarios de una interpretación sintáctica de la relación de indicación no lo harían, porque se dan cuenta de que la relación, sí y cuando exista, no es conceptualmente necesaria: se trata de dos conceptos sin ninguna área de superposición. Por el contrario, si se encuentra una correlación del 0.7 entre el vector de la respuestas a la pregunta ¿cómo juzga las costumbres sexuales de los jóvenes? y el vector de las respuestas a la pregunta ¿cómo juzga una ley que garantice un salario mínimo a cada ciudadano a partir de la mayoría de edad?, es probable que los investigadores perciban un área de superposición semántica entre los dos conceptos (juicio sobre los costumbres sexuales y juicio sobre el así denominado «salario de ciudadanía») y consideren ambos como indicadores del conservadurismo.

Las consecuencias de esta distinción semántica/empírica serán profundizadas en el próximo párrafo, donde veremos que la fuerza probatoria de un coeficiente (es decir, de una herramienta sintáctica) es menor cuando se trata de controlar una relación semántica que cuando se trata de controlar una empírica.

El hecho de que la elección de indicadores sea confiada a una evaluación semántica no significa que la elección sea arbitraria. Obviamente el investigador, o grupo de investigadores, tiene el máximo interés en elegir indicadores que sean válidos (es decir, que le permitan recoger fielmente al menos algunos aspectos de la intensidad del concepto indicado). Por esto no se atreverá a realizar su elección sin conocer la población que va ser objeto de su investigación y la literatura científica sobre el problema que le interesa. Sin embargo, esto no significa que deba elegir sus indicadores en función de su teoría, como sostienen muchos. Sostener que la teoría que se está controlando empíricamente deba dictar, o inspirar, la elección de los indicadores que sirven para controlarla presenta –además de un difuso aire de circularidad– el obvio riesgo de que se elijan los indicadores que ofrecen la mayor probabilidad de corroborarla. Además, el investigador no está nada obligado a repetir las elecciones de los que han estudiado el mismo problema previamente, y todavía menos si estudia el mismo fenómeno en otro ámbito: se asume que el mismo fenómeno pueda tener significados radicalmente diferentes en ámbitos diferentes.¹² Un ejemplo clásico es la tasa de participación electoral, que puede

¹² Esta posibilidad no escapó a muchos autores de comparaciones internacionales (Teune, 1968; Verba, 1969; Frey, 1970; Mokrzycki, 1983), que subrayaban cómo indicadores del mismo concepto tienen que ser «equivalentes», mas no iguales. Sin embargo, esta recomendación no es siempre escuchada, y «muchas instituciones que colaboran con los norteamericanos en investigaciones comparadas se quejan de que los cuestionarios que se les envían desde los Estados Unidos son a menudo acompañados por el pedido de que la traducción sea absolutamente literal» (Brislin, Lonner, & Thorndike, 1973, p. 39).

ser considerada un indicador de fortaleza de la vida democrática en los países donde hay elecciones libres, mientras que era un indicador de la capacidad de movilización del partido único en los países del Este europeo; e incluso podría ser considerada un indicador del cumplimiento de las normas en los países donde votar es un deber jurídico (como en la Argentina).

Otra característica de una investigación que puede cambiar radicalmente el significado de un fenómeno es el nivel de la unidad de análisis, en el caso que esta sea territorial. Un ejemplo de este cambio lo proporciona el diferente significado del porcentaje de ancianos en la población. Si la unidad de análisis es el municipio, esto es un indicador de marginalidad y de estancamiento económico, porque, como se decía antes, los jóvenes tienden a trasladarse a lugares donde encuentran más oportunidades de trabajo. Si la unidad de análisis es más grande e incluye muchos municipios, los municipios periféricos y los centrales se funden en un mismo caso y, por tanto, el porcentaje de ancianos deja de estar vinculado a la diferencia entre municipios centrales y periféricos. A nivel de esta unidad más grande el porcentaje de ancianos puede ser un indicador de calidad del clima y más en general de la vida en las naciones ricas, ya que en muchos países los jubilados tienden a trasladarse, si pueden, a las mejores áreas desde estos puntos de vista.

Más allá de estas consideraciones, si un investigador se aleja de las elecciones de indicadores hechas por los que han estudiado el mismo fenómeno en el mismo ámbito y con (una) unidad de análisis del mismo nivel, él debería justificar su rechazo de los indicadores tradicionalmente elegidos para el concepto que está estudiando y su elección de indicadores diferentes.

5 Los límites de los controles sintácticos de la validez

Por un ridículo complejo de inferioridad respecto a las ciencias físicas, en las ciencias sociales se «mide» todo: desde la provincia de nacimiento al tipo de ocupación; de la religión al estado civil. No se podía imaginar que no se «midiese» la validez, entendida como la relación semántica entre dos conceptos. Esa se mide mediante un singular coeficiente de correlación entre dos vectores: en uno están los datos recogidos mediante el supuesto indicador (sea una pregunta singular, un test, un índice compuesto de varios indicadores) del cual se quiere medir la validez, y el otro representa otro indicador (pregunta, test, índice) cuya validez ya está convencionalmente aceptada por la comunidad científica (*concurrent validation*, Carmines & Zeller, 1979) o bien los resultados de una prueba sucesiva (*predictive validation*, Scott, 1968). El mecanismo de la *concurrent validation* (la más usada) es una patente demostración de la falta de autonomía del control sintáctico: el indicador *I* se considera válido en tanto que estrechamente correlacionado con el indicador *J*, ya bautizado como válido por

la comunidad científica; pero ¿por qué se ha considerado válido el indicador *J*? ¿Debido a su alta correlación con el indicador *Y*? Es evidente que, de remisión a remisión, se abre un regreso al infinito: al final del camino encontramos una evaluación semántica, que los de la visión sintáctica desprecian como no-científica y han bautizado de forma despectiva *face validation* (validación de apariencia).¹³

Más allá de esta inevitable apelación a una evaluación semántica que fundamente el inicio del recorrido, la estrategia de confiar el juicio de validez al juego de los coeficientes de correlación para evitar cualquier intervención de la subjetividad muestra a cada paso su naturaleza ilusoria. Para mostrarlo, en los ejemplos que siguen entraremos en el reino de los coeficientes de correlación y consideraremos cinco variables (*I, J, H, K, Y*) que se quieren evaluar como posibles indicadores del concepto *A*. Los defensores de la visión sintáctica tienen la muy razonable expectativa de que dos o más indicadores del mismo concepto estén positivamente asociados entre sí (Carmines & Zeller, 1979).

Los posibles resultados de lo que acabo de llamar el juego de los coeficientes son innumerables. Empecemos con el menos favorable a las expectativas del investigador «sintáctico»: los diez coeficientes de correlación binarios (en adelante *ccr*) entre los cinco supuestos indicadores resultan bajos. La inevitable conclusión del investigador sintáctico es que los cinco no pueden ser indicadores del mismo concepto –al menos en el área sociocultural donde se realiza la encuesta y en otras áreas afines–. Pero, ¿hay al menos una variable cuyo concepto relativo se pueda adoptar como indicador del concepto *A*? La sintaxis no ofrece ninguna solución a esta pregunta. Para evitar no relevar de ninguna forma el concepto *A* (que nos interesa pero no se puede relevar con una definición operativa directa), la única solución es confiar en el propio conocimiento del contexto de la investigación (es decir, en la semántica) para elegir uno o más indicadores de la canasta inicial.¹⁴

Consideremos otro resultado. Las variables *H* y *K* resultan estrictamente correlacionadas entre sí, pero los otros nueve coeficientes de correlación binarios quedan bajos: ¿esta evidencia de naturaleza sintáctica basta para decidir que *H* y *K* son indicadores válidos de *A*? La fuerte correlación entre *H* y *K*

¹³ También Frey (1970) y Turner (1979) observan que la «validación concomitante» y, en definitiva, todas las formas de validación mediante asociación entre variables en la matriz, abren un regreso *ad infinitum* al final del cual se encuentra –a pesar de la pretendida «objetividad»– la tan despreciada «validación de contenido»: la variable *X* se considera un indicador válido del concepto *C* porque está estrechamente asociada a la variable *Y*; ¿pero quién validó la variable *Y* en cuanto indicador de *C*?

¹⁴ En el §8 (en la segunda parte de este ensayo) se mostrará que nunca es adecuado confiar en un solo indicador. Por tanto, en este caso el dilema es: o renunciar a relevar la propiedad en cuestión, o adoptar como indicadores las dos o tres propiedades que nos parecen semánticamente más cercanas al concepto *A*, a pesar de los bajos *ccr* que tienen entre sí (camino semántico) o las dos o tres variables que tienen los menos bajos *ccr* entre sí (camino sintáctico).

podría tener una naturaleza empírica,¹⁵ o derivar del hecho de que *H* y *K* son ambas indicadores de un concepto diferente de *A*.

Lo mismo puede pasar si las variables estrictamente correlacionadas entre sí son tres o cuatro. Natural y razonablemente, crece en el investigador sintáctico la convicción de haber encontrado tres o cuatro indicadores válidos del concepto *A*. Pero él no puede descartar la eventualidad de que *I*, *J* y *Y*, o *I*, *J*, *H* y *Y* estén estrictamente correlacionadas entre sí porque son todas indicadores de otro concepto, o por ser un efecto artificial de la técnica de recolección de los datos (Campbell & Fiske, 1959; Kriz, 1981, 1988). Y eso puede pasar incluso si las cinco variables elegidas como supuestos indicadores de *A* están correlacionadas.

Otro caso sucede cuando tres supuestos indicadores (digamos *I*, *J* y *Y*) están estrictamente correlacionados entre sí, así como también lo están los dos restantes (digamos *H* y *K*), pero las correlaciones binarias entre cada miembro del primer grupo y cada miembro del segundo son invariablemente bajas. No creo que el investigador sintáctico se conduzca sabiamente si, ahorrando cualquier análisis semántico, decide que –todo considerado– 3 es más que 2. Mi conocimiento de las ideas de colegas me dice que, entre los muchos partidarios de la objetividad, la conciencia de que el resultado de este partido, como de todo otro, depende integralmente de la decisión acerca de cuáles jugadores se ponen en la cancha (en este caso, de cuáles conceptos se consideran como posibles indicadores), una decisión que ninguna computadora, y ninguna rama de la matemática, puede tomar «objetivamente» en lugar del investigador.¹⁶

En resumidas cuentas, incluso una fuerte asociación positiva entre supuestos indicadores no demuestra nada de forma conclusiva: con una feliz expresión de Winch y Campbell (1969), una fuerte asociación positiva es *evidence, not proof* (indicio empírico, no prueba). Solo un adecuado y competente control semántico puede robustecer la convicción de que los resultados numéricos tienen algún sentido más allá de la técnica.

En la segunda parte de este ensayo criticaremos una forma más sofisticada de negación del rol de la semántica y presentaremos la tesis de un doble pluralismo (desde abajo y desde arriba) de la relación de indicación.

¹⁵ Ver arriba (§4) para las diferencias entre relación empírica y relación semántica.

¹⁶ Entre las innumerables manifestaciones de este malentendido y autocontradictorio culto a la objetividad, se puede mencionar el rechazo a eliminar de una canasta de variables sometidas al análisis factorial aquellas que no saturan la dimensión que se está investigando (en el caso de *single-factor analysis*) o ninguna de las dimensiones que emergen del análisis (en el caso de *multiple-factor analysis*). Se considera «no objetiva» cualquier modificación a la canasta inicial, como si esta no fuese a su vez una decisión «subjetiva» del investigador sino algo establecido por una divinidad (Marradi, 1981; Di Franco & Marradi, 2014). Esta muy difundida actitud recuerda la de los judíos que han pasado el Mar Muerto y esperan que Moisés regrese del Monte Sinaí: «Han construido un cordero de oro [nos cuenta el libro *Exodus* de la *Bibla*] y lo adoran».

Referencias

- Blalock, H. M. (1961). *Causal inferences in non-experimental research*. Chapel Hill: University of North Carolina Press.
- Blumer, H. (1948). Public opinion and public opinion polling. *American Sociological Review*, XIII, 542-554.
- Brislin, R. W., Lonner, W. J., & Thorndike, R. M. (1973). *Cross-cultural research methods*. Nueva York: Wiley.
- Campbell, D. T., & Fiske, D. W. (1959, marzo). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, LVI(2), 81-105.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Londres: Sage
- Cronbach, L. J., & Meehl, P. E. (1955, julio). Construct validity in psychological tests. *Psychological Bulletin*, LII(4), 281-302.
- Di Franco, G., & Marradi, A. (2014). *Factor analysis and principal component analysis*. Milán: Franco Angeli.
- Dodd, S. C. (1942). *Dimensions of society*. Londres: MacMillan.
- Durkheim, E. (1893). *De la division du travail social*. París: Alcan.
- Durkheim, E. (1896). *Le suicide*. París: Alcan.
- Frey, F. W. (1970). Cross-cultural survey research in political science. En R. T. Holt & J. E. Turner (Eds.), *The methodology of comparative research* (pp. 173-294). Nueva York: Free Press.
- Galtung, J. (1967). *Theory and methods of social research*. Londres: Allen & Unwin.
- Goffman, E. (1956). *The presentation of self in everyday life*. Edimburgo: University of Edinburgh Social Sciences Research Centre, Monograph 2. (Ed. revisada Nueva York: Doubleday, 1959).
- Guttman, L. A. (1950). The basis for scalogram analysis. En S. Stouffer (Ed.), *Measurement and prediction* (vol. IV, pp. 60-90). Princeton: Princeton University Press.
- Hathaway, S. R., & McKinley, C. J. (1972). Where have we gone wrong? The mystery of the missing progress. En James N. Butcher (Ed.), *Objective personality assessment* (pp. 24-44). Nueva York: Academic Press.
- Kriz, J. (1981). *Methodenkritik empirischer sozialforschung. Eine problemanalyse sozialwissenschaftlicher forschungspraxis*. Stuttgart: Teubner.
- Kriz, J. (1988). *Facts and artifacts in social science. An epistemological and methodological analysis of empirical social science research techniques*. Nueva York: MacGraw-Hill.
- Lazarsfeld, P. F. (1958). Evidence and inference in social research. *Daedalus*, LXXXVII(3), 99-130.
- Lazarsfeld, P. F., & Barton, A. H. (1951). Qualitative measurement in the

social sciences: Classifications, typologies, and indices. En D. Lerner y H. D. Lasswell (Eds.), *The policy sciences. Recent development in scope and method* (pp. 155-92). Stanford: Stanford University Press.

Marradi, A. (1978). Die faktorenanalyse und ihre rolle in der entwicklung und verfeinerung empirisch nützlicher konzepte. *Kölner Zeitschrift für Soziologie*, XXX(3), 488-513.

Marradi, A. (1979). Aggregazione dei comuni in comprensori socio-economicamente omogenei mediante l'analisi fattoriale: il caso della Toscana. *Quaderni dell'Osservatorio elettorale*, (5), 5-54.

Marradi, A. (1981). Factor analysis as an aid in the formation and refinement of empirically useful concepts. En E. F. Borgatta & D. J. Jackson (Eds.), *Factor analysis and measurement in sociological research: A multi-dimensional perspective* (pp. 11-49). Londres: Sage.

Marradi, A. (2002). Método como arte. *Papers. Revista de Sociología*, (67), 107-127.

Marradi, A. (2007). *Metodologia delle scienze sociali*. Bologna: Il Mulino.

McKennell, A. C. (1973, diciembre). Surveying attitude structures: A discussion of principles and procedures. *Quality and Quantity*, VII(2), 203-94.

Merton, R. K. (1948, octubre). The bearing of empirical research upon the development of social theory. *American Sociological Review*, XIII(5), 505-15.

Mokrzycki, E. (1983). *Philosophy of science and sociology. From the methodological doctrine to research practice*. Londres: Routledge & Kegan Paul.

Nicéforo (1921). *Les indices numériques de la civilisation et du progrès*. París: Flammarion.

Nowak, S. (1976). *Understanding and prediction. Essays in the methodology of social and behavioural theories*. Dordrecht: Reidel.

Przeworski, A., & Teune, H. (1970). *The logic of comparative social inquiry*. Nueva York: Wiley.

Quetelet, L. A. J. (1869). *Physique sociale ou Essai sur le développement des facultés de l'homme*. Bruselas: C. Murquaedt.

Reynolds, P. D. (1971). *A primer in theory construction*. Indianápolis: Bobbs-Merrill.

Scott, W. A. (1968) *Attitude Measurement*, pp. 204-273 en Gardner Lindzey y Elliot Aronson (eds.), *Handbook of Social Psychology*, Vol. II. Reading: Addison-Wesley.

Singer, J. D. (1982). Variables, indicators, and data. *Social Science History*, VI(2), 181-217.

Sullivan, J. L., & Feldman, S. (1979). *Multiple indicators: An introduction* (Vol. 15). Newbury Park (Ca): Sage.

Teune, Henry (1968) *Measurement in Comparative Research*, in "Comparative Political Studies" I, 1 (april): 123-38.

Indicators: Syntactic Vision and Semantic Vision: First Part

Thurstone, L. L., & Chave, E. J. (1929). *The measurement of attitude: A psychophysical method and some experiments with a scale for measuring attitude toward the Church*. Chicago: University of Chicago Press.

Verba, S. (1969). The uses of survey research in the study of comparative politics: Issues and strategies. En Stein Rokkan *et al.* (Eds.), *Comparative Survey Analysis* (pp. 56-105). París: Mouton.

Villermé L. R. (1840). *Tableau de l'état physique et moral des ouvriers employés dans les manufactures du coton, de laine et de soie*. París: Renouard.

Winch, R. F., & Campbell, D. T. (1969, mayo). Proof? No. Evidence? Yes. The significance of tests of significance. *American Sociologist*, IV(2), 140-3.