

FRANCIS KORN

QUÉ ES UNA VARIABLE EN LA INVESTIGACIÓN SOCIAL

0.- Introducción

El objeto de este trabajo es definir algunos elementos básicos de la terminología metodológica de las ciencias sociales. Dichos términos (variables, indicadores, categorías, índices) son los nombres de los símbolos que utilizan estas ciencias para presentar a sus referentes empíricos, en el proceso de puesta a prueba de sus enunciados. Como toda ciencia fáctica, la sociología debe trabajar con enunciados susceptibles de comprobación empírica y por tanto debe manejar a sus conceptos de modo que éstos sean pasibles de medición en las pruebas de refutabilidad de los enunciados que los contienen. Para poder decir que las clases sociales tienen condiciones de vida diferenciales, por ejemplo, hay que realizar por lo menos tres operaciones básicas: 1) definir el concepto de clase de modo que sea posible encontrar grupos de individuos que se diferencien significativamente en cuanto a por lo menos uno de los rasgos que establece la definición; 2) definir “condiciones de vida” de modo tal que a cualquier grupo elegido de individuos pueda fijársele un valor que represente este concepto; 3) comparar los grupos de individuos diferenciados significativamente por su posición de clase en cuanto a los valores de “condiciones de vida” que obtuvo cada grupo. En este sentido, puede decirse que una variable es todo concepto capaz de asumir distintos valores.¹

En nuestro ejemplo, clase social y condiciones de vida son variables en la medida en que a cada uno de los dos conceptos se le pueden fijar distintos valores: para clase social los valores podrían ser clase alta, media y baja, o cualquier otro sistema de valores que dependerá del tipo de sociedad donde se quiera realizar la medición y de la precisión que se quiera y se pueda alcanzar con la misma.

¹ En términos de teoría de conjuntos, se puede decir, que la relación que existe entre un conjunto de individuos (Conjunto A) y un conjunto de valores de una variable (Conjunto B) es el siguiente: Si A es el conjunto de objetos a subdividir (por ej. el conjunto de miembros de una sociedad) y B es el conjunto de valores de una variable (por ej. el conjunto de categorías de la variable “clase social”) la clasificación de los miembros de la sociedad según clases sociales será: una función del primer conjunto en el segundo (función de A en B) tal que a cada elemento del primero (A) le corresponda uno y sólo uno del segundo (B). (En nuestro caso: a cada individuo del conjunto A se le asigna una categoría del conjunto B, es decir, cada individuo corresponde a una y sólo una clase social) de modo que la subdivisión del conjunto A se realiza del siguiente modo: Dos elementos del conjunto A están en el mismo subconjunto de A si y sólo si van a parar (por la función) al mismo elemento de B.

1.- Qué es un indicador

El concepto de clase social usado en nuestro ejemplo, si ha de ser parte de un enunciado contrastable será en primer lugar definido conceptualmente (definición real, nominal o de algún otro tipo) de acuerdo al marco teórico general en que el enunciado que lo contiene está inserto. Esta definición puede estar más o menos alejada de los referentes empíricos que realmente van a ser medidos y que representan al concepto objeto de la definición. En general, dadas las características de los conceptos que se manejan en ciencias sociales (tales conceptos generalmente no designan preceptos sino que se refieren a cualidades de objetos y relaciones entre objetos), estas definiciones dejan al concepto relativamente alejado del plano empírico en que va a realizarse la medición. Una definición operacional es aquella que provee de conceptos a nivel empírico, que representan estos otros conceptos. En otros términos, Durkheim enuncia como una regla de su método sociológico que “los hechos sociales no manifiestos han de ser medidos por otros hechos manifiestos que los representan.”²

A estos hechos manifiestos que miden otros no manifiestos, él los llama índices, y es lo que nosotros llamamos indicador. Un indicador de una variable es entonces otra variable que traduce a aquella en el plano empírico. Si la variable que se quiere medir es el grado de alfabetización de una sociedad, un indicador posible será el porcentaje de gente no alfabeto sobre el total de la población de esa sociedad.

Pero puede ocurrir que el grado de complejidad de una variable sea bastante alto, a nivel conceptual, y que su traducción en otra más a nivel empírico sea difícil. El concepto de clase social comporta, generalmente, una serie de aspectos que no están estrechamente vinculados entre sí y que por tanto no pueden representarse por un solo indicador. A estos distintos aspectos de una misma variable (que normalmente son otras variables) las llamamos **dimensiones** de esa variable. Si definimos como la situación de clase de un individuo, el status que éste ocupa dentro de un determinado sistema de estratificación social, y si éste sistema está ordenado según criterios tales como el prestigio ocupacional, el monto de los ingresos familiares, el nivel de educación formal alcanzado y el modo de vida, estos cuatro criterios serán otras tantas dimensiones de lo que llamamos “situación de clase” de cada individuo. Cada una de estas cuatro dimensiones deberá reducirse a indicadores para poder medir concretamente la situación de clase. En este último ejemplo puede verse cómo las distintas dimensiones de la variable clase enunciadas, se sitúan a diferentes distancias del plano empírico: la cantidad de operaciones necesarias para medir “educación formal” es menor que para medir “modo de vida”. Esta última variable debe ser definida conceptualmente y luego

(En nuestro caso: dos individuos caen en el mismo sector de la sociedad si son de la misma clase social).

² Durkheim, E.: “Las Reglas del Método Sociológico”, Bs.As., Dédalo, 1959.

operacionalmente para llegar a un buen indicador de la misma. Para llegar a un valor que represente la situación de clase, según la definición dada, deberemos conocer el valor de cada uno de los indicadores que representa cada dimensión del sistema de estratificación, y componer un índice³ con los cuatro valores obtenidos. Este índice que contiene las cuatro dimensiones del concepto de clase social, podrá ser considerado entonces como un indicador del mismo.

Los pasos hasta ahora vistos para llegar a la medición de una variable, son esquemáticamente los siguientes:

(1)	(2)	(3)	
<u>Variable</u>	<u>Definición conceptual</u> (enumeración de dimensiones)	Definición conceptual de cada dimensión	<u>Definición</u> <u>operacional de</u> cada dimensión. <u>Indicadores</u>

2.- Categorías

Dijimos en el punto 0 que una variable es un concepto capaz de asumir distintos valores. Estos grados o valores que asume una variable son las categorías de esa variable. El sistema de categorías en que subdividimos a una variable: 1) depende de las características de la misma, 2) del contexto de la investigación en que está contenida y 3) del grado de precisión que se quiere alcanzar en la medición. Estas características que parecen casi obvias, es importante recordarlas cuando se elige un determinado sistema de categorías para una variable; es tan importante como recordar los requisitos de cualquier sistema clasificatorio (esto es que las categorías sean mutuamente excluyentes y que el sistema sea inclusivo de todos los elementos que se quieren clasificar)

2.1.- Dijimos en primer lugar que el sistema de categorías que construimos para una variable depende de las características de la misma. Esta proposición sugiere la primera clasificación de variables que podemos enunciar, clasificación que tiene relación con el tema de las categorías ya que subdivide a las variables en tanto a) enumeran magnitudes de una propiedad (o de un concepto) y b) enumeran distintas propiedades de un concepto. Al primer tipo de variables se las llama variables cualitativas o atributos, y al segundo variables cuantitativas o variables propiamente dichas.

³ Conviene distinguir los conceptos de “indicador” e “índice”, que en general se usan como sinónimos, **Indicador** es un concepto más amplio que el de índice, que incluye a este último, y que se aplica a cualquier variable que esté denotando otra variable más alejada del plano empírico. El término “**índice**” se aplica a un indicador complejo compuesto por más de una variable, es decir que contiene más de una dimensión, de la variable denotada.

Así, ingreso, en tanto es subdivisible en intervalos que miden diferencias cuantitativas (sus categorías pueden ser por ejemplo, de \$ 5.000.-; de \$ 5.001.- a \$ 10.000) es una variable propiamente dicha; y participación, en tanto se subdivide en distintas calidades de participación (política, gremial, comunitaria) es un atributo.-

Aquí una aclaración puede ser hecha: en general las variables no son cuantitativas o cualitativas por sí mismas, hasta tanto no se definan dentro del contexto específico de investigación que integran. La variable participación, por ejemplo, si se emplea para determinar la calidad de la participación de un grupo de individuos (en cuyo caso el tipo de categorías sería como el enunciado más arriba) es un atributo: enumera propiedades de un concepto. Pero si con participación se quiere medir el grado en que un individuo interactúa formalmente con los demás miembros de su comunidad, interesa quizás medir la cantidad de asociaciones comunitarias en las que es socio y entonces las categorías de esta variable van a enumerar magnitudes de la misma (el puntaje de cada individuo será 1, 2, 3 etc., su límite superior dependerá de la comunidad donde se realice la medición y las categorías, que también dependerán de este mismo factor, serán del tipo: ninguna, una, dos, tres y más).

La clasificación más general de **escalas de medición** (que no son más que instrumentos precisos para la medición de variables) se basa también en este criterio. Si se sigue la clasificación de escalas de Stevens⁴ que incluye la clasificación de propiedades de un objeto como un tipo de escala, son escalas nominales aquellas que miden un atributo, es decir, ordenan propiedades de un concepto. Aquí, el término ordenación sólo significa asignar símbolos numéricos a cada tipo de propiedad, símbolos cuya asignación es arbitraria ya que el orden puede ser trastocado sin que se altere la clasificación (en nuestro ejemplo de la variable *participación*, considerada como un atributo, de lo mismo asignar el número 1 a cualquiera de las categorías políticas, gremial o comunitaria, y los números 2 y 3 a cualquiera de las restantes).

Las escalas de tipo cuantitativo (ordinales, de intervalos y de razón) son aquellas que miden lo que llamamos una variable propiamente dicha, y que ordenan magnitudes de una propiedad de un objeto. Este tipo de escalas es el que Torgerson considera como escalas de medición⁵, ya que sólo con ellas se cumple el requisito fundamental de medir, que significa asignar números a magnitudes o grados de una propiedad de una serie de objetos.

El sistema de categorías que se construye para una variable cuantitativa depende de la característica de variación de la propiedad que representa esa variable. Ese sistema de categorías debe ser isomórfico con las relaciones entre las varias cantidades (o instancias) en que se subdivide la propiedad que va a ser medida.

⁴ Stevens, : “La Medición y el hombre”, México, F.C.P.

⁵ Torgerson elimina de su clasificación de escalas a la escala nominal por considerar que una mera clasificación no cumple con ninguno de los requisitos que debe llenar una medición.

Las diferencias que pueden establecerse entre variables cuantitativas, en ese sentido, son:

1.- que sólo pueda establecerse el orden de las categorías que representan las distintas instancias. Este es el caso de las escalas ordinales, donde los números asignados a las categorías representan el orden entre las instancias de esa propiedad.

En una clasificación de “prestigio de las ocupaciones”, como ésta se basa en la opinión consensual del prestigio relativo de cada ocupación, la escala resultante es una escala ordinal, ya que todo lo que puede saberse de una ocupación respecto a otra es que tiene más, menos o igual prestigio.

2.- que además de establecerse el orden, pueda establecerse la magnitud de la distancia entre una categoría y otra. Este es el caso de las escalas de intervalos.

Una variable como “edad” o “ingresos” puede subdividirse en categorías de intervalos iguales donde lo que se sabe de un individuo no es sólo que es mayor, menor o igual que otro en edad o en ingresos, sino cuanto es mayor o menor.

2.2.- Decir que el sistema de categorías construido para una variable depende del contexto de la investigación en que está contenida, significa que hay varios sistemas lógicamente válidos para una misma variable. Contexto de la investigación significa no sólo el universo elegido para la misma y el número de individuos que componen la muestra, sino la posición particular de esa variable en el marco de proposiciones construido para ese universo.

Este problema es quizá más fácil de visualizar cuando se trata de construir una escala nominal. Si se trata por ejemplo de realizar una investigación en la que es necesario clasificar a la población inmigrante de Buenos Aires, las clasificaciones de “inmigrantes” en este caso son varias. Si se trata sólo de inmigrantes extranjeros, lo fundamental a tener en cuenta es cuáles son las nacionalidades de inmigrantes, que, por el volumen de los individuos que pertenecen a ella y por la trascendencia que ese grupo nacional tuvo, son pertinentes para ese contexto geográfico y cultural. Si se trata de una investigación que se centra fundamentalmente en el tema de los inmigrantes, el sistema clasificatorio contendrá un número de categorías relativamente amplio, donde exista la posibilidad de discriminación entre varios grupos nacionales más y menos importantes. Si la investigación no se centra en el tema de los inmigrantes, sino que éste es subsidiario, el número de categorías creado en este caso será menor, de modo que permitirá discriminar sólo los grupos nacionales más importantes. En el primer caso, el sistema de categorías contendría, por ejemplo, las categorías de españoles, italianos, franceses, alemanes, ingleses, rusos, otros centroeuropeos, latinoamericanos, asiáticos, africanos, otros extranjeros. La mayor o menor posibilidad de discriminación entre grupos inmigrantes dependería aquí también del número de individuos que constituyen la muestra.

En el segundo caso, cuando el tema de los inmigrantes es subsidiario al objeto central de la investigación y sólo interesa utilizar grupos de inmigrantes como grupos de control respecto a un determinado aspecto de un grupo nativo, las categorías creadas quizá sólo fueran: españoles, italianos, otros europeos, latinoamericanos, otros extranjeros.

Las preguntas importantes a plantearse al construir un sistema de categorías son:

1. Cuáles son los grupos más significativos respecto a la variable que se va a medir;
2. De esos grupos, cuáles son los que interesa discriminar especialmente y comparar con el resto como un todo y
3. Cuál es el número de individuos de la muestra y por lo tanto, cuáles son los grupos que una vez discriminados son suficientemente representativos como para permitir la comparación.-

Respecto a este tercer punto, si el número de individuos de un grupo no es suficiente como para permitir operar con él, lo importante a decidir en cada caso es: a) si la discriminación de ese grupo es indispensable, habrá que elegir una técnica de ponderación o ampliación numérica del mismo; b) si el grupo no es indispensable para la comparación con otros, subsumirlo en una categoría más amplia y heterogénea que funciones como grupo de control.-

2.3.- Que el sistema de categorías depende del grado de precisión que se quiere alcanzar con la medición, significa que el sistema (como vimos en el punto 2.2) puede tener distinto poder discriminatorio dentro del universo a clasificar. Un alto grado de precisión implica generalmente un poder discriminatorio alto. La decisión sobre la precisión de una edición depende de las posibilidades técnicas al alcance del investigador y de la relevancia de la variable a medir, como vimos en el punto anterior.-

De lo visto hasta aquí y para completar el esquema de la página 2, vemos que los pasos para llegar a la medición de una variable se completan con la construcción del sistema de categorías de la variable y de sus indicadores.-

3.- Una clasificación de las variables de acuerdo a la función que cumplen dentro de una proposición:

De acuerdo a la función que cumplen dentro de una proposición, las variables se clasifican en:

- a) Independientes
- b) Dependientes
- c) Intervinientes
- d) Contextuales

Es **independiente** la variable que actúa como determinante en una proposición y **dependiente** la que actúa como resultado, en la proposición:

“El grado de racionalidad política depende del nivel de educación alcanzado”

El “nivel educacional” es el determinante del “grado de racionalidad”, que por tanto, es su resultado. Así, la primera, “nivel educacional”, es la *variable independiente* (porque los valores que tome no dependen de la otra variable con que se la relaciona en la proposición) y la segunda, “racionalidad política” es la *variable dependiente* (porque los valores que tome dependen de la variación de la primera variable).-

Es importante notar aquí, que ninguna variable es dependiente o independiente por sí misma. Su clasificación en este sentido depende, como ya dijimos, de su posición dentro de una proposición. Lo que sí es cierto es que en cierto tipo de proposiciones, que relacionen un variable de tipo estructural (con este nombre queremos denotar variables que se refieren por ejemplo a la estructura económica o demográfica de la sociedad global) o en general una característica permanente con una variable de tipo actitudinal, lo normal es que la primera variable funcione como variable independiente con respecto a la variable actitudinal, porque es difícil suponer lo contrario. Si queremos por ejemplo, establecer alguna relación entre la procedencia rural urbana y el grado de prejuicio de una población, seguramente la “procedencia” funcionará allí como el determinante del grado de prejuicio.-

Una variable interviniente es, como su nombre lo indica, una variable que “interviene” en la relación entre otras dos. “Intervenir” significa aquí, especificar, explicar, interpretar o decidir que la relación entre las dos variables iniciales es espúrea. En el ejemplo de Durkheim sobre la relación entre la tendencia al suicidio y la pertenencia a distintos grupos religiosos, una de las variables intervinientes que él usa es la “cohesión del grupo”. En primer lugar él encuentra que hay relación entre la tendencia al suicidio y la religión que se profesa, ya que en diferentes contextos los protestantes se suicidan más que los católicos, luego explica esta mayor propensión al suicidio de los protestantes introduciendo la variable “cohesión del grupo”, de donde los protestantes se suicidan más porque tienen menor cohesión grupal.-

Aquí llamaremos de modo general, “interviniente” a toda variable que se introduzca en la relación entre otras dos, sea esta variable temporalmente “antecedente” (es decir es una característica del grupo temporalmente anterior a la que se considera independiente) o realmente interviniente (es el caso del ejemplo de Durkheim citado: una variable que es contemporánea de la independiente).-

Llamaremos “contextual” a una subclase de las variables intervinientes. Se designarán así aquellas variables que especifiquen para qué ámbito geográfico, social, cultural, se establece la relación entre dos variables. Si decimos: “En Buenos Aires las clases altas son

más prejuiciosas que las bajas”, el término “Buenos Aires”, se puede designar como un valor de la variable contextual “ciudad”. Otros valores de esa misma variable contextual podrían ser otras ciudades que se usasen para la comparación.-

Quizás a esta altura no quede demasiado claro por qué dijimos que a las variables contextuales las consideramos una subclase de las intervinientes. Esto es así porque la función que cumplen las variables contextuales es la misma que la que cumple cualquier variable interviniente de características distintas (actitudinales, estructurales, etc.) Es decir, sus funciones son también de especificación, explicación, interpretación o anulación de la relación anterior.-