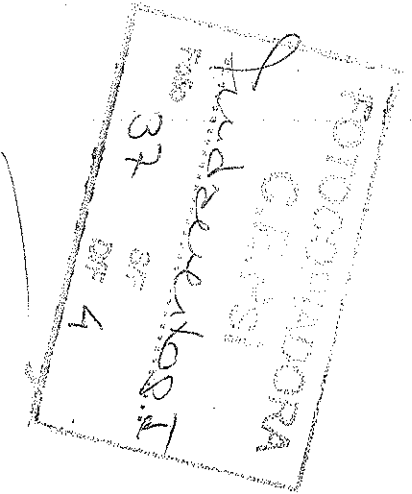
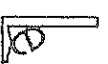


NURIA CORTADA DE KOHAN

# Teorías Psicométricas y Construcción de Tests



 LUGAR EDITORIAL

## CAPÍTULO I

# Posibilidad de medición en Psicología

*"La ciencia moderna al formar la mente humana hacia un análisis de los hechos exacto e imparcial, es una educación especialmente adecuada para promover una cabal ciudadanía".*

KARL PEARSON  
*"La gramática de la ciencia", 1911.*

EN CUALQUIER TEXTO DE PSICOLOGÍA GENERAL se suele explicar que un test psicológico es un instrumento de medición para estudiar algún aspecto de la conducta de las personas, tales como la inteligencia, las aptitudes, el rendimiento escolar, la personalidad, los intereses o las actitudes sociales. Pero ahí radica de inmediato lo primero que debemos plantearnos desde el punto de vista científico: ¿Es posible medir la conducta humana? ¿Es posible la medición en psicología? Esta es una pregunta crucial pues hace a la existencia misma de la psicología como ciencia. En efecto, el principal objetivo de una ciencia más que la mera descripción de los fenómenos empíricos, es establecer mediante leyes y teorías, los principios generales con los cuales se pueden explicar y pronosticar los fenómenos empíricos. Para llevar a cabo este objetivo las ciencias se dedican a recoger y comparar datos para obtener las correlaciones y expresar funcionalmente con expresiones matemáticas, las teorías que son el fin de su búsqueda. En gran parte, la medición permite que el lenguaje matemático se aplique a la ciencia y ya dijo con gran visión Galileo que "la naturaleza es un libro abierto para quienes saben usar el lenguaje matemático".

## DIFERENCIA ENTRE LA EXPLICACIÓN CIENTÍFICA Y EL SENTIDO COMÚN

El hombre siempre se ha preguntado: ¿Porqué actúo de este modo? ¿Porqué unas personas son diferentes de las otras? Los antiguos miraban al cielo en busca de respuestas usando analogías superficiales. Así, si un niño nacía durante la constelación de Leo creían que estaba destinado a ser fuerte y altivo como un león. Más tarde buscaron explicaciones de la conducta en la experiencia externa: quien se pareciera a un zorro debía ser astuto, y el que se parecía a un lobo, voraz. Así la astrología y la fisiognomía son dos caminos por los que el hombre ha tratado de explicar su conducta. La respuesta del sentido común a los problemas del conocimiento del hombre han pasado de generación en generación en forma de proverbios y sabiduría popular. Al mismo tiempo, muchos filósofos han meditado sobre la naturaleza humana y han construido grandes sistemas, a menudo contradictorios entre sí. Sin duda, muchas de sus concepciones poseían parte de verdad. Pero la mayor debilidad de todas las explicaciones precientíficas es que no ofrecían ningún método para dilucidar las contradicciones y así nos encontramos con proverbios que postulan la bondad y certeza de principios absolutamente contradictorios. Los filósofos no cometían errores cuando trabajaban los problemas que implicaban tan solo la deducción de premisas que son verdaderas. Los errores surgían cuando se suponía que ciertas premisas eran verdaderas y no lo eran, construyendo sobre ellas un débil castillo de cartas. Poco a poco, el hombre fue descubriendo que para conocer la naturaleza era necesaria la investigación científica. Primero surgieron las ciencias formales como la lógica y la matemática, luego las ciencias fácticas naturales como la física y la química. Desde hace unos trescientos años la ciencia ha sido cada vez más importante para el mundo biológico y en los últimos cien años ha comenzado a dedicarse científicamente al estudio psicoló-

gico del hombre y al estudio social de las organizaciones humanas. (Kerlinger, 1988; Lazardfeld, 1955)

Las explicaciones científicas son diferentes de las del sentido común porque los fundamentos sobre los que se aceptan o se rechazan son distintos. Acá la fuerza de la tradición, el encanto de la analogía, la erudición o el sentido creador de un autor no son suficientes. En la ciencia la autoridad proviene solamente de la evidencia de los hechos derivada de la observación sistemática y repetible (Bunge, 1972). A medida que surgen nuevos hechos se comprende que las ideas aceptadas durante siglos deben rechazarse o ser revisadas. En cambio, conceptos aparentemente extraños o inverosímiles son examinados y aceptados como lo fueron en su momento las ideas de Copérnico cuando se comprendió el error de Ptolomeo. La ciencia es *investigación* y esto implica que siempre busca, prueba, revisa y reforma. Investigar proviene de la voz latina *vestigium* que literalmente es "planta del pie" y por extensión la "huella" que queda o sus indicios. Los sinónimos de investigar son: indagar, inquirir, examinar, inspeccionar, explorar, buscar y rastrear. Buscar precisamente, coincide con el sentido de la palabra inglesa *research*. La palabra *research*, y este proviene del francés antiguo *sercher* (actualmente *chercher*) que a su vez deriva del latín *circare*, esto es, rodear. Así "research" sería propiamente una búsqueda intensiva de indicios, efectuando un rodeo. Es obvio: se investiga lo que no puede aprehenderse de un modo inmediato. El proceso de una investigación siempre refleja este modo de llegar al conocimiento de algo con esfuerzo, por la vía indirecta de un "rodeo" siguiendo una huella, un vestigio o indicio por un largo camino en forma sistemática, es decir, con "método". El supuesto básico sobre el que se construye la ciencia es que esta es un universo regulado por leyes en el que existe coherencia entre los hechos. La ciencia trata de relaciones coherentes. El psicólogo, por ejemplo, considera que entre las distintas conductas hay una coherencia explicable por leyes generales e intenta hallar estas leyes que han de permiti-

tirle comprender la conducta de los seres humanos. (Arnau Grass, 1990; Northrop, 1947).

El conocimiento de la coherencia de los hechos nos permite realizar mejores predicciones sobre la conducta de las que haríamos sobre la base de la intuición, adivinando o haciendo conjeturas al azar y analogías superficiales y además somos capaces de estimar los límites de predicciones a través de la observación de la variabilidad de los hechos. (Szilassi, 1956; Hempel, 1956; Towsend, 1953).

En psicología científica solemos hablar de predictores y no de "causas". El uso de la palabra causa supone una relación en la cual un hecho (la causa) invariablemente precede y produce otro (el efecto). Pero en psicología la mayoría de los hechos suelen ser el resultado de varias causas y por ello es mejor hablar de "predictores" que de causas. (Brown y Ghiselli, 1956).

El investigador se enfrenta con problemas. En forma general podemos decir que existe un problema cuando no encontramos en forma inmediata y directa respuesta para una pregunta. Si ahora volvemos a analizar en que se diferencia el punto de vista ingeniero del sentido común, con el enfoque científico vemos que, aunque los métodos científicos surgieron del sentido común, en el proceso de su evolución se produjeron cambios fundamentales: 1) el sentido común se organizó sistemáticamente y 2) se modificó el orden en que trabajaba. Por ejemplo: supongamos un fenómeno psicológico bien conocido como el de que los recuerdos se desvanecen a medida que pasa el tiempo. Cuando el hombre común se enfrenta con esto pretende explicarlo diciendo: "cuando no se usa lo que se aprende, por falta de uso su recuerdo se pierde". Pero los científicos como Ebbinghaus ya en 1885 no se conformaron con esta respuesta simple. Trabajando metódicamente con sílabas sin sentido quisieron comprobar si era solamente por falta de uso que los hechos se olvidan y entonces hallaron que lo que provoca la disminución del recuerdo es principalmente lo que se llamó la *inhibición reactiva* que es el efecto de nuevos aprendizajes durante el

intervalo entre el aprendizaje original de lo que nos interesa y el del recuerdo. La diferencia entre los dos enfoques es, entonces, que con el sentido común quedamos satisfechos con una explicación aparente, mientras que el científico busca sistematizar los hechos para sobrepasar la explicación obvia.

El principal objetivo de una ciencia, más que la mera descripción de los fenómenos empíricos es establecer leyes y teorías, los principios generales con los cuales explicar y pronosticar los fenómenos. Al llevar a cabo esta finalidad, la ciencias se dedican a recoger y comparar datos para obtener las relaciones y las ecuaciones matemáticas que han de permitirles establecer las hipótesis, leyes y teorías que explican un campo de la realidad. La medición es una de las cosas que hace posible estos procesos pues la medición hace que el lenguaje matemático se aplique a la realidad estudiada. (Cohen y Nagel, 1934).

#### ESTRUCTURA DE LAS CIENCIAS

Una ciencia consiste por un lado, en teoría y por otro, en datos, fenómenos o evidencias empíricas. La parte teórica son constructos ideales, el aspecto empírico son datos observables. El científico vincula ambos estableciendo reglas de correspondencia, que sirven al propósito de definir ciertos constructos teóricos en términos de datos observables y estas reglas de correspondencia son en última instancia los procesos de medición. Cualquier investigación suele comenzar con una idea formulada como hipótesis que tiene el investigador de un proceso. Por ejemplo, supongamos que un teórico nos dice "que cuando menor es el nivel social de una persona, menor será el grado de conocimientos que posee". Esta es una *definición teórica o constitutiva*. Para poder demostrar la verdad o falsedad de esta hipótesis el investigador la debería expresar con la *definición operacional* correspondiente, es decir se debería definir "nivel social" determinando con qué escala se ha evaluado nivel social y deberíamos también señalar el test o examen especial con que hemos medi-

do el "grado de conocimientos" de una persona. Con una *definición operacional* todas las personas que usan el procedimiento logran resultados similares, mientras que las definiciones teóricas tienen siempre un alto grado de subjetividad.

### NATURALEZA DE LA MEDICIÓN

Vemos que en las definiciones operacionales que se necesitan para la investigación, la medición de las variables es un problema fundamental. La medición es un concepto básico en todas las ciencias y es más difícil de captar en las ciencias sociales y del comportamiento que en las ciencias físicas y naturales. La definición de medición que nos parece más adecuada siguiendo a Torgerson es la de Campbell (1938) quien sostiene que "medición es la asignación de números para representar las propiedades de los sistemas materiales en virtud de las leyes que gobiernan estas propiedades". Un sistema material es un objeto; por ejemplo un árbol. Una propiedad es una *variable*, por ejemplo altura, peso, dureza, diámetro del tronco, etc. La medición de una variable indica, por lo tanto, una relación de *isomorfismo*, es decir de uno a uno, una equivalencia de formas, entre la estructura lógica del sistema numérico y la estructura de la naturaleza que se manifiesta en las variables o propiedades que se miden.

### CARACTERÍSTICAS DE LOS NÚMEROS

Fundamentalmente son tres: *Orden, distancia y origen*.

1) *Orden* (Significa que los números están ordenados de mayor a menor:  $1 < 2 < 3$ , etc.)

2) *Distancia* (Las diferencias entre los números también están ordenadas)

Por ejemplo:

$$8 - 5 > 7 - 5$$

$$10 - 8 < 10 - 7$$

$$5 - 3 = 7 - 5$$

3) *Origen* (La serie numérica tiene un origen único que llamamos "cero" y la diferencia entre cualquier par de números que tiene a 0 en un miembro es igual al número del otro miembro.)

$$5 - 0 = 5$$

La medición en Psicología ha sido difícil de aceptar en parte por la gran influencia de dos grandes pensadores como Kant y Bergson que habían dicho que las matemáticas no podían aplicarse a la Psicología. Sin embargo, la medición es posible porque la estructura del pensamiento del hombre y la actividad psicológica en general poseen propiedades des- de el punto de vista lógico que son suficientemente similares a la estructura de las matemáticas. Es posible, por lo tanto, establecer un isomorfismo. Sin embargo las discusiones sobre medición siguen. Por ejemplo, Lord y Novick (1968) definen la medición como un "procedimiento para la asignación de números (puntuajes o medidas) a propiedades específicas de unidades experimentales de modo que las características y preserven las relaciones especificadas en el dominio comportamental". Las reglas en el sentido de Stevens y el "preservar las relaciones" de Lord y Novick suponen que para representar la propiedad debe existir un isomorfismo entre las características del sistema numérico y las relaciones entre las diversas cantidades de la propiedad medida. El problema de la construcción de escalas ha recibido una gran atención en la psicología desde los trabajos de Stevens siendo actualmente la *Teoría Representacional de la medición* la posición más ortodoxa en cuanto a la conceptualización de la medida. Esta teoría es axiomática y formalizada y trata el tema de la medición articulándola en tres grandes áreas: el problema de la *representación*, *el de la unicidad* y *el de la significación*. La teoría tiene su origen en los trabajos de Hölder y Russell alrededor de 1900; pero quienes han dado formulaciones más completas son Luce, Krantz, Tversky, Roberts y Mitchell (1990). No podemos entrar en detalle en estas nuevas teorías que suelen estudiarse en Psicología Matemática. Digamos sólo que desde el punto de vista de la *representación* la medición supone encontrar un sistema relacional numérico con una estructura semejante al relacional empírico que se pretende medir. Dada esta semejanza uno de los sistemas puede utilizarse para representar al otro. El problema de la *unicidad* hace referencia a la arbitrariedad de los números

Si aceptamos el criterio de Stevens (1951) que es el más amplio, podemos dividir los niveles de medición en cuatro escalas fundamentales, nominal, ordinal, de intervalos iguales y de cocientes.

#### a) Escala nominal

Es el nivel más primitivo. Según Coombs (1981) y Torgerson no debería considerarse un nivel de medición, pues no mide variables, solo clasifica atributos. La asignación de números es arbitraria. Por ejemplo, la clasificación de pacientes psiquiátricos, los números de las licencias de los autos, la clasificación de profesionales, etc. La única condición que requiere es que los miembros o elementos de una clase o categoría deben ser equivalentes, idénticos, respecto al atributo que tenemos en cuenta.

Solo acepta los postulados lógicos de los números de identidad o equivalencia.

#### b) Escala ordinal o por orden jerárquico

Cuando los elementos de una categoría no sólo son distintos de los de otra sino que entre ellos hay un orden que se expresa con los símbolos "mayor que" o "menor que" (" $>$ " o " $<$ ").

Por ejemplo, la clasificación de los sujetos en analfabetos, con educación primaria, secundaria o universitaria. Acepta pues los postulados de los números de orden jerárquico además de los de identidad.

#### c) Escala de intervalos iguales

Acá las distancias numéricamente representan distancias iguales en la variable empírica que miden. Por ejemplo, medimos la temperatura con la escala de tipo intervalos iguales. Se supone que la diferencia entre 30 grados y 35 es la misma que entre 80 y 85. Sin embargo, no hay cero absoluto que corresponda con la nada de la propiedad medida, por esto el cero de la escala centígrados es igual a 32 grados en la escala Fahrenheit. El 0 se ha colocado en la congelación del agua, pero esto es arbitrario, se hubiera podido colocar en la

elegidos según la teoría representacional. Una vez establecidas las relaciones numéricas es posible asignar diferentes conjuntos de números a los elementos del sistema manteniendo el homomorfismo, es decir, pueden obtenerse distintas escalas de números para la misma variable o atributo. El problema de la *significación* se refiere a la validez de una conclusión empírica inferida a partir de una conclusión numérica. Esta validez siempre es relativa al tipo de escalas en que se basan las inferencias. Stevens plantea la solución del problema en términos estadísticos admisibles para cada tipo de escala.

Un planteo distinto al de la teoría representacional es el de la teoría de la medición conjunta de Luce y Tuckey (1964) que también es una teoría axiomática. Los procedimientos clásicos de cuantificación (medición extensiva) suponen encontrar una forma de combinar las cantidades empíricas que refleje directamente la naturaleza cuantitativa de la variable. La medición conjunta permite detectar la estructura cuantitativa de una variable a través de relaciones ordinales observadas entre sus valores.

### NIVELES DE MEDICIÓN

No todas las variables son perfectamente isomórficas con los números, por lo tanto las variables se pueden medir con mayor o menor perfección según la exactitud con que ellas se corresponden con las propiedades lógicas de los números. (Cortada de Kohan, 1994).

En este sentido de la información que proporcionan los números, tenemos distintos tipos de escalas o niveles de medición. Sin entrar en las discusiones sobre definiciones, muy bien aclaradas por Torgerson en "*Theory and methods of Scaling*" (1958), digamos que existen tres criterios: según Russell "la medición de magnitudes es cualquier método por el cual se establece una correspondencia única y recíproca de todas o algunas magnitudes de una clase y todos o algunos números". Para Campbell medición es "la asignación de números para representar propiedades de los sistemas materiales" y para Stevens "medir es asignar números a los objetos o hechos de acuerdo con reglas conocidas".

congelación del aceite o de cualquier otra sustancia. Casi todas las escalas de tipo psicológico usadas en los tests para medir inteligencia o aptitudes se considera que son así. Sus transformaciones son del tipo  $y = ax + b$ , para  $a > 0$ .

d) *Escala de razones o cocientes*

Las escalas de mediciones físicas, en cambio, son las más perfectas pues no sólo tienen intervalos sino que suele tener un cero absoluto, es decir el cero de la escala numérica coincide con la nada de la propiedad a medir como las escalas de longitud, peso, densidad, etc. y se llaman de cocientes porque acá no sólo existe la misma relación entre intervalos o diferencias sino que hay la misma relación entre dos puntos de una escala medida por distintos sistemas. Por ejemplo, con la escala de pesos en kilogramos y en libras:

|    |   |     |     |     |     |    |      |
|----|---|-----|-----|-----|-----|----|------|
| Kg | 0 | 1   | 2   | 3   | 4   | 5  |      |
| l  | 0 | 2,2 | 4,4 | 6,6 | 8,8 | 11 | etc. |

tenemos que :

$$(3 - 2) / (6,6 - 4,4) = 0,454$$

$$(5 - 4) / (11 - 8,8) = 0,454$$

pero en el cociente entre dos puntos también hay igualdad, es decir, existe igualdad entre el cociente de dos puntos cualquiera, esto es:  $3/2 = 1,5$  y  $6,6/4,4 = 1,5$  y esto, como puede averiguarse fácilmente el lector interesado, no ocurre comparando las escalas de temperatura en grados centígrados y en Fahrenheit.

Esto tiene importancia, porque las escalas de medición por cocientes permiten muchas más operaciones estadísticas que las escalas de intervalos iguales y su ecuación transformadora es del tipo  $y = a \cdot x$ , para  $a > 0$ .

CLASES DE MEDICIÓN

La medición en ciencia responde, además, a una categorización como la siguiente, enunciada por Campbell, que de mayor a menor precisión sería:

1) *La medición fundamental o medición de magnitudes A.* Acá los números se asignan según leyes naturales que representan la propiedad misma. Por ejemplo: la longitud, el volumen, la resistencia física poseen significado constitutivo y no presuponen la medición de ninguna otra variable.

2) *También tenemos la medición derivada o medición de magnitudes B.* Acá la característica de una escala obtiene sentido mediante leyes que relacionan una propiedad con otras propiedades. Así la densidad de una sustancia es la relación entre su masa y su volumen. Esta relación difiere en las distintas sustancias y por lo tanto el valor de esta relación puede medir la densidad de la sustancia.

3) Finalmente está la medición por *fiat*. El significado acá se debe a una definición arbitraria que depende de relaciones supuestas entre las observaciones y el concepto. Esto es lo común en las ciencias sociales cuando usamos índices, por ejemplo, el nivel socioeconómico o en psicología el cociente intelectual o la capacidad de aprender de un animal por cantidad de ensayos que necesita para aprender a recorrer un laberinto o solucionar un problema. Los índices tienen solamente significado operacional, no constitutivo.

*Enfoque para la construcción de escalas*

Es importante ahora que comprendamos que todos los instrumentos de medición psicológicos conocidos, tests y pruebas mentales, escalas de actitudes, tests de resolución de problemas, inventarios de personalidad, pruebas de conocimiento o técnicas de proyectivas, etc., siempre presentan estímulos a los que esperamos que los *sujetos* den una respuesta.

Siempre se trata de una escala de medición (hecho de uno o varios estímulos) que se aplica a uno o varios sujetos para obtener una o varias respuestas. Estas tres palabras, *estímulos*, *respuestas* y *sujetos*, son importantes para conocer la naturaleza del instrumento. Tenemos tres enfoques posibles:

1) *El enfoque centrado en el sujeto*

La mayor parte del campo de la psicometría esta basada en mediciones que atribuyen la variabilidad de las respu-

tas a las diferencias individuales de los sujetos. El propósito de estos tests es escalar a los sujetos gracias a una serie de estímulos o ítem que se consideran réplicas. Así, si tenemos un test de conocimiento con 50 preguntas tenemos alumnos que contestarán bien 25, 30 o 50 ítem y podremos ordenarlos desde los que tienen menos conocimientos hasta los que tienen más. El esquema corresponde a un diseño de análisis de variancia simple. La variación sistemática de los resultados se atribuye a diferencias entre los sujetos.

### 2) Enfoque centrado en los estímulos

En estos tests, de los que hay pocos, como las escalas de actitudes por intervalos aparentemente iguales de Thurstone, la variación sistemática se atribuye a la diferencia de los estímulos respecto a un determinado atributo. Los sujetos son considerados réplicas, por esto se les llama "jueces".

### 3) Finalmente el enfoque centrado en las respuestas,

poco utilizado, considera que la variabilidad de las reacciones frente a los estímulos puede atribuirse a dos fuentes de variación, la de los sujetos y la de los estímulos. El esquema teórico sería el de un análisis de variancia doble por hileras y por columnas con una sola observación por celda. Se puede ejemplificar este enfoque con la técnica de Guttman para el análisis de actitudes.

## MEDICIÓN Y EVALUACIÓN

En psicología y en educación no todos los instrumentos tienen el mismo rigor científico, en su elaboración. Veremos en el capítulo VI que elaborar un instrumento de medición no es tarea fácil, pero en general todos los instrumentos conocidos tienen el propósito de evaluar alguna característica de los sujetos. Cuando este propósito de dar una valoración ordinal de alguna variable se alcanza pero sin pretensiones de llegar a una medición más exacta hablamos de evaluación, como por ejemplo los exámenes orales o escritos. Con el fin de abarcar un ámbito amplio a nuestra tarea es que la llamamos de evaluación y medición.

## BIBLIOGRAFÍA

- ARNAU GRASS, JAIME (1990) *Diseño experimental en psicología y educación*. México Ed. Trillas
- BUNGE, M. (1972) ¿Qué es un problema científico? *Cuadernos de Epistemología*. Buenos Aires, Fac de Filosofía y Letras. Publicación interna.
- BROWN, C.W. y GHISELLI, E.E.(1956) *Scientific Method in Psychology*. Nueva York. Mc Graw Hill.
- CAMPBELL, N.R. (1952). *What is science?* Nueva York, Dover.
- COHEN, M.R. y NAGEL, E. (1934) *An Introduction to Logic and Science Method*. Nueva York. Harcourt and Brace.
- COOMBS C. H. DAWES R.M. y TVERSKY, A. (1981) *Introducción a la Psicología Matemática*, Madrid, Alianza.
- CORTADA DE KOHAN, NURIA. (1994) *Diseño estadístico*. Buenos Aires, Ed. Eudeba.
- HEMPFL, C.G. y OPPENHEIM, P.(1956) *Fundamentals of Concept Formation in Empirical Science*. Chicago, Univ. Press.
- KERLINGER, F.N. (1988) *Foundations of Behavioral Research*. Holt, Rinehart, Wiston.
- LAZARSFELD, S. y ROSENBERG, R. (1955) *The Language of Social Research*. Illinois, The Free Press Glencoe.
- LORD, F. M. Y NOVICK M.R. (1968) *Statistical Theories of Mental Test Scores*. Reading Mass. Addison and Wesley.