

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO.

FACULTAD DE

PSICOLOGIA.

"MEDICION EN PSICOLOGIA."

TESINA.

para obtener el título de
LICENCIADO EN PSICOLOGIA.

P R E S E N T A .

RAUL TENORIO RAMIREZ.

México, D. F.

1973.

DO CUBIR
ESTAS HOJAS



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mis Padres y

Hermanos : Aurelio,

Alberto, Vicente, Carmen

y Ma. Luisa.

A mis Maestros que con su ayuda,
logre este paso crucial en mi vi
da académica.

Luz Ma. Barraza.

Ma. Luisa Preciado.

Alfonso Andrade.

Con un gran afecto al amigo
que con sus consejos logre-
llegar a este peldaño de la
vida.

Prof. José J. Medrano P.

A mis maestros que me
ayudaron en mi forma-
ción profesional.

Dr. Rogelio Díaz Guerrero.

Mtra. Ma. Luisa Morales C.

Dr. José Cueli García.

A la mujer que con su ayuda
material y moral me impulsó
en la vida académica y per-
sonal.

ALICE.

I N D I C E .

	Página
INTRODUCCION.	1
I.- ANTECEDENTES.	
a) La Medición en Ciencia.	3
b) Definiciones de los que es la Medición.	5
c) La Medición en Psicología.	9
II.- LA CONSTRUCCION DE MODELOS DE MEDICION.	
a) La Medición como base de la Construcción de Modelos.	16
b) Evaluación de los Modelos.	18
c) Modelos de Escalas para la- Medición de Respuestas.	21
d) Modelos de Escalas para la- Medición de Estímulos.	25
e) Modelos de Escalas para la- Medición de Personas.	27
BIBLIOGRAFIA.	37

I N T R O D U C C I O N .

El hombre en su afan de saber "que tanto hay de este y de aquello", desde tiempos muy remotos ha querido medir lo que está a su alcance y también lo que tiene -- más allá.

Remotamente, quizá conocer medidas de distancia tales como que tantos metros había entre una ciudad y otra; pero después no solo le interesó el lugar donde se -- desarrollaba sino también el espacio sideral.

Y con su afan lo logro, queriendo actualmente medir todo lo que se encuentra en su medio ambiente y -- así su interés por medir creció hasta llegar a lo que ahora trataremos en este trabajo "midio al hombre en sus capacidades", su "inteligencia", su "personalidad", etc.

Pero existe un simple detalle, que para hacer estas mediciones no tenía instrumentos con que hacerlo y tuvo que empezar a construirlos tomando ejemplos de otras ciencias; como la Física; y con el tiempo los ha ido refinando y ha construido nuevos y mejores métodos e instrumentos de medición para "el hombre"; cosa que aquí discutiremos muy brevemente.

Describiremos lo que es la medición y su importancia en la ciencia y específicamente en Psicología y -- finalmente los modelos que se han construido para efectuar-

dicha medición.

I.- ANTECEDENTES.

a).- La Medición en Ciencia.

El principal objetivo de una ciencia es la --- descripción empírica de fenómenos, es establecer a través de leyes y teorías, principios generales por medio de los cuales el fenómeno puede ser explicado empíricamente, medido y pre-- dictivo.

Varias disciplinas científicas difieren unas -- de otras en variedad de maneras, una clasificación sugerida - por Margenau ésta de acuerdo al grado de procedimientos teó-- réticos ó de que explicaciones son usadas como contrastadas - con procedimientos correlativos ó explicaciones.

La estructura de una ciencia fué dibujada como- una interrelación compleja entre la teoría y los datos; la -- teoría consiste en el trabajo de construcciones y sus interco_o necciones y los datos se obtienen de los fenómenos observa--- bles del mundo.

Las conecciones formales entre las construccio- nes en una manera y las reglas de correspondencia en otra for_o ma establecen dos clases diferentes de definiciones usadas en ciencia. La definición formal es llamada constitutiva y más - tarde se dan definiciones operacionales ó epistémicas.

Las construcciones útiles en ciencia deben te-- ner importancia sistemática y empírica. Por importancia siste_o mática entenderemos que la construcción debe ser tal que nos-

permita la formación de múltiples conexiones con otras -----
construcciones en la estructura. Por importancia empírica en-
tendemos que la construcción debe ser conectada directamente-
ó a través de otras construcciones a los datos observables. -
Una cosa importante para la ciencia es la investigación para-
la invención de construcciones con importancia empírica y sis-
temática. Las construcciones científicas son de varias clases.
Los sistemas incluyen los objetos y cosas de experiencia dia-
ria. Las propiedades se refieren a los aspectos observables -
y la medición a las propiedades de sistemas y no a los mismos
sistemas. La medición fué vista como una manera de definir --
una propiedad.

Las construcciones usadas en ciencia pueden ser
clasificadas dentro de varios tipos diferentes. Para nuestros
propósitos presentes solo sería necesario distinguir entre --
sistemas y propiedades.

Las propiedades son esencialmente los aspectos-
ó características observables del mundo empírico, por lo que-
siempre que definamos una propiedad, siempre parecerá ser una
propiedad de algo. Para este algo usaremos el término sistema.
En este caso, las propiedades, donde ocurren, son aspectos ó-
características de los sistemas.

Mientras, está distinción entre sistemas y sus-
propiedades es tal vez obvia, sin embargo, es una distinción-
importante; ya que el hecho es que siempre las propiedades --
son medidas y no los sistemas. Medir es siempre la medición -
de una propiedad y nunca la medición de un sistema.

b).- Definiciones de lo que es la medición.

Debido a la naturaleza de la medición parece -- difícil definir el término. Russell dice que "Medición de magnitudes es; en un sentido más general; algún método por el -- cual una correspondencia única y recíproca es establecida entre todas ó algunas de las magnitudes de una clase y todos ó algunos de los números integrantes, racionales ó reales según sea el caso".

Campbell dice que medir "es la asignación de números a propiedades representantes de materiales de otros sistemas, en virtud de leyes que gobiernan estas propiedades".

S. S. Stevens (1951) quien ha trabajado en ésta área ha señalado, que en un amplio sentido, medir es "el proceso de asignar números a objetos de acuerdo a ciertas reglas". Exactamente que reglas son usadas en una situación particular dependen del tipo de comparaciones que puedan ser hechas con los objetos a los cuales los números serán asignados.

Estrictamente hablando no se permite usar las reglas que siguen aquellas manipulaciones con números que no pueden ser bien manejados por comparación de los mismos objetos. Existen cuatro grupos de reglas, generalmente uno de los cuatro grupos de reglas es usado. Medir con estos diferentes grupos de reglas se refiere a medir en cuatro diferentes tipos de escalas, cada una de las cuales representa un nivel diferente de cuantificación de la variable estudiada.

La forma más simple de cuantificación involucra nada más la asignación de niveles numéricos a objetos indivi-

duales ó clases de objetos y esto se refiere a medir en una escala nominal. Un ejemplo de este tipo de medida sería la asignación arbitraria de números a varios grupos ocupacionales. La regla solo observada, es que a todos los individuos de la misma ocupación les será asignado el mismo número y no a dos individuos que tengan diferentes ocupaciones se les dará el mismo número. Cuando esta cuantificación simple es trabajada, el único tipo de operación matemática que permite es el conteo; por ejemplo : determinando el número ó proporción de individuos que tiene cada ocupación. De ésta manera una escala nominal nos permite establecer diferentes clases de objetos.

En el otro extremo, medir en una escala de razón, donde los valores asignados a objetos se conducen parecidamente a números reales y entonces puede ser usada apropiadamente con todas las operaciones matemáticas, excepto aquellas que son exclusivas a los números imaginarios. Este tipo de medidafrecuentemente disponible en las ciencias físicas, nunca es logrado en la medición de rasgos conductuales. Las escalas de razón requieren de un cero absoluto (esto es, cero en la escala debe representar completa ausencia de la característica involucrada), así como también unidades iguales de medidaa través de la escala (por ejemplo : las diferencias entre 2-y 3 en una escala deben representar el mismo grado de diferencia en la característica que se ésta midiendo, como una diferencia entre los valores 16 y 17 de la escala).

Entre estos niveles extremos de cuantificación-

se encuentran los dos tipos de escalas más frecuentemente usa dos en la medición psicológica. La escala intervalar y la ordinal. La principal diferencia entre la escala de razón y la intervalar, es que la primera demanda un cero absoluto, sin embargo, las dos requieren de igualdad en sus unidades a lo largo de la escala.

Tal vez la mejor manera de ilustrar las limitaciones impuestas, es decir, cuando debe ser usada una escala de intervalo ó una ordinal más que una de razón, se muestra en la analogía de como el resultado sería parecido si las res tricciones tomadas son semejantes en una característica familiar como lo es la estatura. Medir en una escala de intervalo ésta variable sería como si fuera necesario medir las estaturas de un grupo de individuos tomando como base un cero arbitrario (tomándolo como base para la elaboración de una tabla). Así se puede ver que la persona A tiene una estatura de 1.60-cms. en la tabla que reporta las estaturas y que la persona B mide 1.50 cms. en la misma tabla. Ciertamente esto haría posi ble decir que la persona A, para quien la distancia desde la base de la tabla a donde termina su cabeza fué 1.60 cms. es más alta que la persona B; para quien la misma distancia fué solo 1.50 cms. Y con ésta sola restricción no podrían decirse las diferencias que existen entre las alturas de estos dos -- individuos. Y no podría tampoco decirse que la persona que mi dió 1.60 cms. de estatura en la tabla es el doble de alta que la persona que midió 1.50 cms. de altura en la tabla (Si la tabla empieza con 30 cms. de altura, las alturas actuales se-

rían de 1.90 y 1.80 cms. respectivamente). Así cuando la medida esta hecha sin un cero absoluto (esto solo se puede hacer en escalas de intervalo) no es legitimo, multiplicar ó dividir los valores registrados como medidas. Aunque estas operaciones aritméticas pudieran ser ejecutadas por supuesto que los propios números no lo permitirían. Y la gran dificultad que se presenta, es que algunas de las inferencias y conclusiones hechas como resultado de tales procedimientos serían engañosas. Afortunadamente las escalas de medición de intervalos iguales son muy usadas en la evaluación de características conductuales, y ellas si permiten el manejo de la mayoría de los índices estadísticos (tales como la media, la desviación estandard y la correlación) los cuales son usados en esta área. El problema real es que la mayoría de los datos crudos registrados como una descripción numérica de un rasgo conductual rara vez son logrados en este nivel de medición. La mayoría de las calificaciones tienen solo las propiedades de una escala ordinal.

Lo análogo de la medida de la estatura con el cero arbitrario como base de la tabla puede ser un cuidadoso y amplio ejemplo para ilustrar las restricciones impuestas -- cuando se tienen datos que solo se pueden manejar con las operaciones matemáticas que admite la escala ordinal.

Esto nos da una idea de que, sin embargo, las condiciones restringidas, descritas en la variable altura son impuestas sobre la mayoría de las medidas de características conductuales, esto es el uso de muchos procedimientos ma-

temáticos comunes o la interpretación directa de calificaciones pueden llevarnos a serios errores. Tomando en cuenta estas circunstancias tenemos tres alternativas.

La primera alternativa es simplemente restringirse uno solo, al uso de aquellos procedimientos estadísticos desarrollados especialmente para el uso con datos intervalares, tales como la media, la T de Kendall y otras. La segunda posibilidad es que asumiendo (el porque de alguna evidencia empírica ó deducción lógica) que el rango fundamental esta distribuido normalmente en la población y usando los procedimientos descritos por Guilford (1954) convertir la información ordinal a una escala intervalar. Con muchos de los procedimientos más comunes, sin embargo, se reconoce el hecho de que medir características conductuales es manejar cantidades relativas y no absolutas. Esto significa que los valores "crudos" originalmente registrados nunca serán interpretados directamente. Mejor dicho debe hacerse la conversión a alguna de las formas de calificaciones "derivadas", que son las que relacionan la ejecución de un individuo a un grupo total de personas. Entonces y solo entonces por la definición exacta y cuidadosa que se hace de un punto cero y una unidad de medida que se considera que es igual a través de la escala, pueden usarse los procedimientos matemáticos usuales y hacer las interpretaciones sin conceptos falsos.

c).- La Medición en Psicología.

La medición de las cantidades físicas tales como el peso, la estatura y el tiempo tienen una escala con va-

lor de cero significativo y esto puede ser razonable para ---
asumir que el intervalo entre puntos iguales en varios luga--
res a lo largo de la escala representa igualdad acerca de la-
cantidad que está siendo medida. Desafortunadamente esto no -
puede hacerse en Psicología. Una calificación de 0 en una ---
prueba de "inteligencia" no significa una ausencia total de -
habilidad psicológica. No puede decirse que ganarse una cali-
ficación de 10 a 20 reactivos correctos en alguna prueba de -
rendimiento (logro) representa el mismo incremento en ejecu--
ción que si se gana de 70 a 80 reactivos correctos. Por lo --
tanto los psicólogos, más que otros investigadores de otras--
áreas, han trabajado arduamente con las bases lógicas de la -
medición.

"Conocer" a una persona significa realmente ser
capaz de describirla precisa y plenamente, pero no podemos --
describir a una persona en todo lo que atañe, por lo cual de-
bemos de elegir un atributo de acuerdo a lo que nos interesa-
en el momento. Solo podríamos describirla con precisión y en-
términos cuantitativos, en cuanto más querámos esto, más ten-
dremos que recurrir a las mediciones que constan de tres pa--
sos comunes y que son los siguientes :

- 1.- Señalar y definir la cualidad ó atributo --
que se habrá de medir.
- 2.- Determinar un conjunto de operaciones en --
virtud de las cuales el atributo pueda mani-
festarse y hacerse perceptible.
- 3.- Establecer un conjunto de procedimientos ó-
definiciones para traducir las observacio--

nes a enunciados cuantitativos de grado ó de cantidad.

La comprensión y los problemas que presentan -- estos pasos nos proporcionan el cimiento sólido para la comprensión de los procedimientos y de los problemas de la medición en psicología.

Primero, nunca podremos medir cosas o personas, lo que mediremos siempre será un atributo de estos. Siempre mediremos estatura, temperatura, gusto, "inteligencia", madurez emocional, etc. En el establecimiento de estos atributos tan sensibles nos olvidamos de la definición ó significado -- del atributo. Al hablar de los términos largo y corto sabemos a lo que nos estamos refiriendo, pero no siempre sucede lo -- mismo con todos los atributos físicos. Hasta no llegar a definir claramente el atributo no podremos medirlo con sentido. -- En la medida en que no estemos de acuerdo en lo que significa el atributo, en esa medida discerniremos acerca de cuales son los mejores procedimientos para medirlo. El problema es mayor cuando pensamos en los atributos que manejan los psicólogos, y debido a su naturaleza, es más difícil llegar a una -- definición clara y precisa de estos. Qué entendemos por "inteligencia"? Cuáles serán las conductas que calificaremos como inteligentes ? . Todos entenderíamos lo que queremos decir -- con conducta inteligente en un sentido general pero al ahondar más en el problema las diferencias de conceptualización -- se harían más grandes, lo mismo puede decirse de casi todos los conceptos psicológicos, de unos más que de otros. Siendo-

el primer problema al que nos enfrentamos, el definir claramente el atributo que vamos a medir. Junto con este problema tenemos el de que atributos debemos medir primero y conforme a este tendremos que hacer la descripción del objeto o persona, de ésta última, la descripción tendrá poco valor si elegimos rasgos que no vengan al caso.

El segundo aspecto de la medición es el de encontrar ó crear un conjunto de operaciones que aislarán el atributo en el que estamos interesados y lo pondrán ante nuestros ojos.

Las operaciones para medir la longitud de un objeto se encontraron en remotos tiempos de la historia de la humanidad. En la escuela se les enseña a los niños la regla, la cinta de medir, el metro; que son los instrumentos aceptados universalmente; y el solo hecho de colocarlos junto al objeto que se quiere medir se considera propio para poner de manifiesto la longitud de este. Pero cómo mediremos las distancias entre una ciudad y otra ? , entre el sol y la tierra ? , entre una nebulosa y otra ? , ó como medir la longitud de un bacilo de tuberculosis ? , ó el diametro de una neurona ? .La Física ha creado los instrumentos y las respuestas necesarias para realizar esto, ya sean cantidades demasiado grandes ó muy pequeñas. Las operaciones para medir la longitud ó la distancia se han vuelto indirectas, complejas y cada vez más precisas. Y se les emplea porque nos proporcionan resultados mejores, más consistentes, verificables y útiles.

Existe una influencia recíproca entre la defi--

nición de atributo y las operaciones necesarias para ponerlo de manifiesto. La definición que elegimos determinará que es lo que aceptaremos como operaciones pertinentes o razonables al mismo tiempo que, las operaciones que ejecutamos ponen de manifiesto el atributo ó sea la definición de éste. A esto lo llamaremos "definición operacional". Lo que estamos diciendo es que el procedimiento que en nuestra opinión mostrará el atributo se convertirá en la definición efectiva por lo que a nosotros respecta.

La historia de la medición psicológica se ha -- realizado durante el presente siglo, se ha realizado la creación de instrumentos y procedimientos para poner de manifiesto, de manera estandarizada y en condiciones uniformes, las conductas que sirven de indicadores de los atributos de las personas. Por ejemplo : el trabajo que ha realizado Binet proporciona definiciones operacionales de "la inteligencia".

Existe mucha ambigüedad en las definiciones por una parte, y una gran variedad en los instrumentos que hemos creado para poner de manifiesto las conductas que queremos medir y por la otra la jerarquización de las personas de diversas maneras con respecto a las diferentes medidas. Una vez -- que se ha definido el rasgo se nos presenta la siguiente pregunta : Cuáles serán los procedimientos u operaciones que podemos utilizar para estimar su presencia ó ausencia ? .

Otro problema que existe al respecto, es la --- elección del mejor instrumento para este fin. En qué se debe fundar para hacer esta elección ó estimación ? . Existen tres

consideraciones específicas que dividiremos en tres categorías.

Que son : Validez, Confiabilidad y Valor Práctico.

Por validez entenderemos el grado en que nuestro instrumento mide lo que queremos medir realmente.

La confiabilidad es la exactitud y precisión de un procedimiento de medición.

El valor práctico tiene que ver con una gran variedad de factores de : economía, conveniencia y posibilidad de interpretación.

La función esencial de la medición es la agrupación en orden, de una clase de eventos con respecto a la exhibición de una propiedad particular y esto transmite el descubrimiento de una clase ordenada de elementos que pueden ser colocados en correlación de uno a uno con los eventos que se están tomando en cuenta.

La razón fundamental por la que debemos establecer una teoría de la medición psicológica es la naturaleza del universo de observación en psicología. Otra razón será la de que las habilidades o rasgos que los psicólogos desean estudiar en general no son directamente medibles, deben de ser estudiados a través de otras propiedades.

Si un profesor elabora dos exámenes finales, una persona se examina con los dos, obteniendo diferentes calificaciones en los dos exámenes. En general el orden de las

calificaciones en un grupo para el primer exámen no será el mismo que para el segundo exámen. Estas diferencias en las calificaciones representan un error del instrumento de medición. Cuando deseamos medir la habilidad de una persona con dos instrumentos perfectamente contruidos, pueden dar una misma o semejante calificación para el examinado, en tanto su habilidad no cambie, estas diferencias son ocasionadas por errores de medición. Esto solo funciona con las ciencias físicas pero no así con las ciencias conductuales. En física se puede medir repetidas veces lo mismo.

En psicología se podría repetir la misma medida una ó dos veces pero no podríamos hacerlo más, por los cambios que tuviera la persona debido a la fatiga ó al aprendizaje. El físico desea obtener una medición segura de un objeto, y en Psicología se prueba a un individuo en un grupo y se busca la relación con el grupo y la evaluación del mismo sujeto.

Como ciencia, la Psicología está unificada y organizada a través de teorías psicológicas. Una teoría psicológica es un sistema de afirmaciones relativas a grupos de conceptos, los cuales les sirven para predecir y explicar algunos aspectos limitados del dominio de la conducta. Generalmente una teoría bien desarrollada contiene uno ó más modelos formales que dan estructura concreta a los conceptos generales de la teoría, estos modelos pueden ser considerados como explicaciones parciales de la teoría general, y están relacionados directamente con fenómenos observables. La función de los modelos es permitir la deducción lógica de relaciones

generales y específicas, que no han sido demostradas empíricamente. Ninguna teoría científica es hasta ahora considerada - descriptora de la verdad, pero la asociación de los modelos-- es considerada más ó menos útil en varios contextos.

II.- LA CONSTRUCCION DE MODELOS DE MEDICION.

a).- La Medición como Base de la Construcción-- de Modelos.

La medición en Psicología ó en cualquier otra - ciencia empieza con un procedimiento para identificar los ele^umentos del mundo real con los elementos ó construcciones de - un sistema lógico abstracto (modelos) a través de la defini^ución semántica precisa de los elementos básicos de la teoría. En los últimos años, sin embargo, la psicología a venido ---- usando la medición y realmente es una sistematización de la - medición en forma de modelos psicológicos; los cuales son básicamente en forma matemática.

Las construcciones de la teoría psicológica se expresan en palabras. Estas palabras son mucho más explíci--- tas y definidas, más restrictivas, sintácticas y semánticas-- que cuando se usan en su significado más amplio en el lenguaje diario ó en teoría precientífica. Un modelo matemático difiere de uno puramente verbal en algunos aspectos. Primero, - es identificado con un sistema matemático exacto generalmente de muy alto orden (álgebra ó cálculo), por medio del cual --- las construcciones elementales pueden ser manipuladas para fa^ucilitar deducciones a partir del modelo. También aunque el mo^udelo matemático es más preciso, evita la confusión que se pro

duce por la imprecisión de las afirmaciones del modelo paramente verbal. Sin embargo, debido a que las construcciones del sistema son más explícitas definidas, tienden a tener menor significado connotativo y así, tal vez menor riqueza conceptual. Finalmente, el modelo matemático atrae y representa solamente aspectos más limitados del dominio de la conducta. Los modelos de rasgos latentes, han sido ofrecidos en parte, como un medio para conectar los modelos matemáticos de la conducta más precisos aunque limitados con las teorías psicológicas más ampliamente concebidas.

A pesar de la regla mencionada anteriormente para medir, puede ser tomada como una medida si ayuda a tener cierto plan internamente consistente para el desarrollo de la nueva medida. El plan es conocido como modelo de escala y la medida que resulta de ejecutar el plan es mencionada como una escala. Los métodos para construir y aplicar reglas constituyen el modelo de escala. El propósito de cualquier modelo de escala es generar un continuo en el cual se localizan personas u objetos.

En cualquier problema particular de medidas, la escala se refiere a una matriz de datos (tabla) que se obtiene de una serie de "estímulos" dados a "personas" y éstas a su vez emiten "respuestas".

Cuando se observan cuidadosamente los diferentes modelos para escalar se verá que en casi todos los casos, no está presente una de las dimensiones de la matriz de datos, por ejemplo : está solo una persona, ó un solo estímulo,

ó solo un tipo de respuesta a cada estímulo.

Si los datos no se ajustan a las suposiciones de un modelo en particular para la escala unidimensional, el investigador tiene tres caminos a seguir :

- 1.- Probar uno de los modelos para la escala unidimensional.
- 2.- Poner a prueba métodos de análisis multivariados ó .
- 3.- Poner a prueba algún otro problema.

b).- Evaluación de los Modelos.

Con frecuencia existen diferentes modelos que pueden ser usados para el desarrollo de ciertas escalas y en algunas ocasiones los modelos conducen a diferentes conclusiones sobre las propiedades de la escala de los datos. Un modelo podría rechazar los datos por estar de acuerdo con una escala ordinal, mientras que otro modelo podría aceptar los datos por estar de acuerdo con una escala de intervalo. Cómo podemos escoger el modelo apropiado para nuestro problema particular ? . No existe una forma segura de conocer esto por anticipado. La prueba definitiva y final es lo bien que encajan las escalas derivadas en un nexo de relaciones legales con otras variables. Para esto pueden aplicarse ciertos criterios de sentido común, parte de esto se refiere a la atracción intuitiva de un modelo de escala. A pesar de que los datos científicos deben ser objetivos, el científico debe confiar en su

intuición para sus ideas de investigación.

Otro aspecto del "sentido común" es la selección del modelo de escalas y se refiere a la evaluación de las suposiciones en los modelos en términos de lo que ya se conoce sobre el tipo de datos de que se trata. Después de que un modelo se usa para derivar una escala y antes de que se hagan esfuerzos vigorosos para encontrar sus relaciones legales con otras variables, tenemos ciertas formas preliminares de evidencia referente a lo útil de la escala. Si los valores de la escala para las personas ó los objetos se ven afectados ligeramente por la forma de recopilar los datos seguramente la escala no funcionará bien en la práctica. Un tipo más importante, aún de evidencia preliminar se refiere al monto del error de la medición; implicado en el uso de un tipo particular de escala. Una escala con un alto grado de error no puede ser utilizada en ninguna forma para ningún propósito.

Anteriormente se dijo que los problemas de las escalas se refieren a una matriz tridimensional que contiene a personas, a estímulos y a respuestas.

En una escala unidimensional por lo general una persona tiene un solo tipo de respuesta a cada estímulo. En cualquier modelo se aplica una matriz bidimensional. Pero lo que debería aclararse son los métodos usuales de escalar que se utilizan para los estímulos que son diferentes a los empleados para las personas y también debe aclararse cual de las dos variables que van a ser escaladas tienen una mayor influencia sobre la forma de obtener las respuestas. Sería

más fácil pensar en los problemas de medición en términos de escalar a las personas.

Tenemos una cosa más fina, cuando el objeto es escalar estímulos y no personas. Cuando se trata de escalar unidimensionalmente los estímulos se espera encontrar una escala que se amolde al modelo individual. Por lo tanto una escala de este tipo sería típica de las personas, como grupo, a pesar de que no pudiera representar perfectamente la escala que se puede obtener mediante una intensiva investigación de cierta persona. El propósito de las investigaciones a largo plazo es escalar estímulos y no personas ó sea relacionar las escalas de los mismos estímulos con respecto a diferentes atributos. Por mucha importancia que tenga el escalamiento de estímulos en Psicología, lo más importante será escalar personas con respecto a los atributos.

En la literatura psicológica encontramos que -- las investigaciones que se han hecho se han realizado sobre escalamiento de personas y no de estímulos; los estudios que se han trabajado han sido en aprendizaje, ansiedad, toma de decisión, "inteligencia" y el vigor de las respuestas condicionadas; todo esto relacionado definitivamente con personas, ya sea que estos estudios sean para ver diferencias individuales ó para experimentos controlados.

Es importante hacer una distinción entre el escalamiento de estímulos y el escalamiento de personas, debido a que hay un mayor número de problemas con los primeros. En este este escalamiento con frecuencia el tema de las investi-

gaciones se refiere a la naturaleza exacta de las relaciones funcionales entre el escalamiento de los estímulos y sus diferentes circunstancias.

La exigencia más importante es que las diferentes escalas de las personas esten monotónicamente relacionadas unas con otras, ó sea que clasifiquen a las personas en la misma forma.

Debido a que hay más problemas con las escalas de estímulos, que con las de personas, la mayor parte de los puntos relacionados con el escalamiento y con la mayor parte de los modelos de escalas, se han presentado en el contexto de los problemas relacionados con la escala de estímulos. Esto lo podemos observar en Guilford (1) y Torjerson (10), - en estos tratados la mayor parte concierne a modelos de estímulos.

Esta diferencia ha tenido una influencia en el lenguaje empleado para describir la investigación psicológica. Cuando hablamos de "Escalar" ó de "Métodos de Escalar" -- por lo general discutimos el problema relacionado con el escalamiento de estímulos; pero cuando discutimos el problema relacionado con el escalamiento de personas en relación a un atributo, es más probable que se empleen términos como "Medidas" y "Construcción de Pruebas" ..

c).- Modelos de Escalas para la Medición de Respuestas.

Antes de discutir los modelos para escalar, es-

nesario revisar algunos de los diferentes tipos de respuestas que se requieren de los sujetos. El tipo de respuesta tiende a correlacionarse con el tipo de estímulo que ésta siendo estudiado. Por lo general se requieren respuestas diferentes para la escala de los estímulos y la escala de las personas. Tenemos también que los diferentes modelos para escalar con frecuencia requieren diferentes tipos de respuestas; aquí discutiremos los tres tipos más importantes de distinciones.

"Juicios y Sentimientos", aunque no existe una definición para estos términos en la teoría de la medición,-- una distinción importante es la que se refiere a respuestas-- relacionadas con juicios y sentimientos. Juicio es cuando hay una respuesta correcta, existiendo una cierta comparación verídica para la respuesta del sujeto.

La palabra sentimiento es para cubrir todas las respuestas relacionadas con las reacciones, preferencias, intereses, actitudes y cosas similares, personales. La diferencia importante entre juicio y sentimiento es que con los sentimientos no existe comparación verídica.

El problema de los juicios es relacionar la intensidad percibida de cierto atributo con su intensidad física. En las ocasiones en que los sujetos hacen juicios existe la comparación verídica. Actualmente se discute mucho acerca de las respuestas correctas.

"Respuestas comparativas y absolutas" es otra distinción importante del tipo de respuestas que puede dar el sujeto. La primera es cuando el sujeto da una respuesta abso-

luta a cada estímulo y una respuesta comparativa es cuando -- se pide al sujeto que indique que cosa le gusta más.

Hay muy pocos casos en los cuales tiene senti-- do exigir a los sujetos respuestas absolutas, pues no hay for ma de comunicar su juicio absoluto sobre propiedades físicas.

Resumiendo las respuestas absolutas son útiles-- en algunas situaciones y nos sirven para :

1.- Proporcionar un método breve de lograr in-- formación comparativa y.

2.- Para obtener una indicación aproximada del-- punto neutral en un continuo de sentimientos.

Otra importante distinción entre los tipos de -- respuestas, se refiere a la escala sobre la cual el sujeto de -- be responder. En la mayoría de los casos se le pide que res-- ponda en una escala ordinal, de intervalo ó proporcional ó -- sea que es requisito que él conteste en el tipo de alguna de -- éstas dos escalas. Hay muchas formas de obtener las respues-- tas con respecto a los tres tipos de escala y la forma se lla -- ma Método Psicofísico.

Se requiere que la mayoría de los sujetos fun-- cionen en una escala ordinal y los métodos aplicados se llama -- rán de "Estimación Ordinal", la forma más directa es hacerlo -- por orden de rangos, en el cual el sujeto clasificará de ma-- yor a menor con respecto a un atributo de juicio ó sentimien-- to. Un efecto más completo lo tenemos con el método de compa -- ración de iguales en el cual se le pide al sujeto que clasifi -- que los estímulos, dos al mismo tiempo en todas los pares po --

sibles.

Otro método ordinal es el del Estímulo Constante que es similar al de comparaciones por pares con la excepción de que un estímulo estandar es emparejado sucesivamente con cada miembro de un conjunto constante de estímulos.

Otro método ordinal es el de Categorías Sucesivas, en el cual al sujeto se le pide que seleccione una colección de estímulos de cierto número de categorías que se ordenan de acuerdo a un atributo específico.

Existen muchas variantes del método de Categorías sucesivas, dependiendo del tipo de información ordinal. A pesar de que hay numerosas variantes para estos métodos, los enfoques básicos para obtener las respuestas a juicios son apoyadas en las escalas ordinales y las expresiones de sentimiento lo son en las escalas de orden por rango ó intervalares, comparaciones por pares, estímulos constantes y categorías sucesivas.

Uno de los métodos con mayor frecuencia empleado para obtener respuestas de intervalo por parte de los sujetos es el de los intervalos aparentemente iguales, que es parecido al de categorías sucesivas, la diferencia ésta en las instrucciones.

El método de intervalos aparentemente iguales es de gran utilidad para obtener respuestas de intervalo en una amplia categoría de métodos que mencionaremos como los métodos del cálculo de intervalos.

En particular el método más utilizado en este grupo es el de Bisección, que es un enfoque diferente, el cual le presenta al sujeto dos estímulos que son extremos con respecto al atributo y se le pide que forme un juicio sobre la proporción de los intervalos formados con cada uno de los estímulos insertados en ellos.

En los métodos de cálculo intervalar, se le pide al sujeto que calcule los tamaños comparativos de los intervalos entre los estímulos. Los métodos del cálculo de la proporción requieren que los sujetos respondan a las magnitudes absolutas de los estímulos.

d).- Modelos de Escalas para la Medición de Estímulos.

Después de haberse obtenido la respuesta por medio de algún método, el siguiente paso es generar una escala ordinal, intervalar ó de razón; para poder escalar a los estímulos. En la escala de estímulos no se requieren modelos complejos, ya que generalmente para derivar escalas ordinales en los diferentes modelos usados para éstos fines, se llega al mismo orden de estímulos. Los modelos cobran importancia cuando se trata de construir una escala intervalar ó de proporción para estímulos.

En las escalas basadas en estimaciones subjetivas se requiere que el sujeto responda en términos de una es-

cala ordinal, intervalar ó de proporción. Aún cuando al sujeto se le instruye para que sus respuestas sean de un tipo de escala, el experimentador puede tomar las respuestas como representando otro tipo de escala. O sea el sujeto puede responder en términos de una escala intervalar y el experimentador puede tomarlas en términos de escala ordinal. Al desarrollar escalas con modelos basados en la estimación subjetiva, a veces se asume, bien ó mal ; que el sujeto estaba empleando una forma de escala superior a la que se le indicó que usará.

Existen varios modelos para desarrollar escalas de intervalos y de proporciones, basados en estimaciones subjetivas. La esencia de ellos, es que se supone, que los sujetos son capaces de producir tales escalas directamente.

Después de haberse presentado ésta suposición--fundamental y de emplear los métodos adecuados para la recopilación de las respuestas, especialmente aquellas estimaciones de escalas intervalares y de proporción. Los conjuntos de suposiciones especiales, adicionales permiten la derivación de las escalas intervalares y a veces las de proporción. Los modelos y procedimientos de computación actuales para estos fines tienden a ser más sencillos que los modelos basados en -- otras concepciones fundamentales.

Los modelos discriminantes son la segunda clase importante, de ellos solo dos han alcanzado mayor desarrollo--hasta hoy. Son diferentes en varios aspectos importantes a -- los de la clase anterior que están basados en estimaciones --

subjetivas. En los modelos discriminantes no se toma muy en serio la habilidad del sujeto para generar una escala de intervalo ó de proporción, más bien supone que el investigador lo tendrá que hacer después de recopilar los datos. Estos modelos hacen énfasis en la variabilidad de respuestas por diferentes personas al mismo estímulo y la variabilidad potencial de las respuestas de la misma persona al mismo estímulo en diferentes ocasiones. Mientras que los modelos basados en estimaciones subjetivas exigen respuestas por diversos métodos de estimación de intervalos y de proporción, los modelos discriminantes básicamente exigen respuestas por métodos de estimación ordinal, como el orden por rangos y la comparación por pares.

Los fundamentos teóricos para todos los modelos discriminantes en el escalamiento de los estímulos fueron por L. Thurstone. Para cada individuo al que se le presenta un estímulo existe un proceso discriminativo con respecto a un atributo específico. Un proceso discriminativo es simplemente una reacción ampliamente definida de cierta clase que se correlaciona con la intensidad del estímulo, sobre una escala de intervalos de un atributo existiendo una distribución discriminativa para cada estímulo.

e).- Modelos de Escalas para la Medición de Personas.

El primer modelo es el de medidas multireactivos. La palabra reactivo será empleada en un amplio sentido--

para representar cualquier estímulo usado en los métodos de medición. Por lo tanto los reactivos pueden ser palabras en una prueba de deletrear, comparaciones entre pesos, declaraciones sobre actitudes hacia las Naciones Unidas, las elecciones correctas de una rata en un laberinto y las reacciones en un estudio de tiempos de reacción.

Existen ciertas razones de importancia para exigir más de un reactivo en casi todas las medidas de los atributos psicológicos.

Primero, los reactivos individuales generalmente son muy específicos. Es decir, cada reactivo tiende a tener solo una baja correlación con el atributo que se está midiendo y tiende a relacionarse con otros atributos.

Aún en el caso de que los reactivos individuales no fueran específicos existen otras razones para que las dimensiones requieran más de un reactivo. Una de éstas razones es que casi todos los reactivos tratan de categorizar a las personas, esto puede ser en dos grupos ó en un número más pequeño de grupos. En la mayor parte de los problemas de medición es conveniente efectuar diferenciaciones finas entre las personas y eso rara vez puede lograrse midiendo nosotros con un solo reactivo un atributo.

Aún en el caso de que no exista la especificación entre los reactivos y estos fuesen muy sutiles y capaces de elaborar distinciones muy finas entre las personas quedaría una razón más importante para que no fuera suficiente la medición con un reactivo. Los reactivos individuales tienen -

considerable error que ocasiona una cantidad de error al azar y no puede uno confiarse del reactivo para proporcionar una medición exacta del atributo.

Al combinar reactivos se puede llegar a distinciones muy sutiles entre personas y así la confiabilidad tiende a aumentar cuando el error de medición se reduce, según va aumentando el número de reactivos en una combinación de reactivos. Para ello casi todas las mediciones de atributos psicológicos son multireactivos.

Casi todos los modelos para escalar personas pueden ser representados por diferentes tipos de curvas relacionando un atributo a la probabilidad de responder en una forma u otra ante los reactivos. Las mediciones totalmente confiables del atributo se llaman resultados verdaderos y las aproximaciones a los resultados verdaderos se llaman resultados falsos.

Al discutir las curvas es conveniente pensar en el atributo como completamente continuo, es decir, es teóricamente posible hacer discriminaciones infinitamente pequeñas entre las personas. También es útil pensar que existe un gran número de personas en cada uno de los puntos infinitos del atributo. En ésta circunstancia hipotética la curva muestra las respuestas esperadas de las personas a cada nivel del atributo y lo que se espera se expresa ya sea en posibilidades de respuestas alfa para los reactivos dicótomos ó como un resultado promedio de reactivos multipuntos. Por la naturaleza de las probabilidades van acompañadas de algún error.

Entre los modelos que existen para la medición de personas están los determinísticos, estos se llaman así -- porque se supone que no existe error en las curvas. El modelo determinístico particular más utilizado es aquel que supone -- que hasta cierto punto del atributo, la probabilidad de respuesta es cero y mayor que ese punto la probabilidad de respuesta es uno.

Cada reactivo tiene una correlación biserial de uno con el atributo y en consecuencia cada reactivo discrimina perfectamente en un punto particular del atributo. Intuitivamente este es un modelo muy atractivo debido a que es exactamente lo que uno espera obtener en mediciones de extensión.

En las escalas de Guttman, aunque uno nunca conoce la naturaleza exacta de las curvas se pueden observar -- los datos y ver si se desarrollan en una norma triangular y -- al hacerlo uno ésta haciendo una sutil y lógica suposición.

Para desarrollar esta escala lo que uno hace es administrar un conjunto de reactivos a un grupo de personas y después tratar de arreglar las respuestas en tal forma que -- produzcan la norma triangular, como es poco probable que se -- obtenga la forma triangular exacta y es necesario por lo tanto :

- 1.- Descartar algunos reactivos.
- 2.- Encontrar la mejor forma de ordenar los --- reactivo y a las personas.

En esta última consideración se debe tener principal consideración con la reproductibilidad de las normas de

resultados.

Si se obtiene la forma triangular perfecta, el conocimiento del número de las respuestas contestadas si por el sujeto permite a uno reproducir todas las respuestas de la persona. A diferencia de cuando se obtiene aproximadamente la forma triangular el conocimiento de las respuestas contestadas si de una persona permite una reproducción aproximada de las respuestas de esa misma persona.

La importancia de la escala de Guttman es que se puede investigar el porcentaje de reproductibilidad de todas las personas y todos los reactivos.

Aparte de la escala de Guttman existen otros modelos determinísticos, entre los cuales está el no monótono que hace las siguientes suposiciones. Cada reactivo recibe -- respuestas alfa por todas las personas de cierto nivel y cada persona responde en la forma alfa a un solo reactivo. En contraste con la escala de Guttman en este modelo determinístico cada reactivo tiene una curva no monotónica; es decir, la línea sube y después baja.

Empleando el término "reactivo" en su sentido más amplio, sería muy raro encontrar algún reactivo sobre medidas psicológicas que se ajustará a este modelo.

Los modelos determinísticos son de aplicación -- principalmente para los especialistas en la teoría de la medición psicológica. Tales modelos con frecuencia representan -- "casos limitados" de modelos que actualmente se usan para desarrollar medidas de los atributos psicológicos. Únicamente tra

bajando con algún tipo de modelo no determinístico de probabilidades se pueden desarrollar las medidas que se necesitan en la investigación.

Otro de los modelos es el de probabilidad para la escala de personas. Si nosotros supieramos que las curvas no tuvieran subidas perpendiculares y bajadas semejantes, se podrá trabajar con algún tipo de modelo de probabilidad. --- Existen numerosos tipos de modelos de probabilidades, dependiendo del tipo de curva asumida por las líneas de trazo. A continuación veremos algunos de estos modelos :

Modelos no Monótonos, que son los de probabilidades y son análogos a los modelos no monótonos determinísticos como el discutido anteriormente. Cualquier tipo de curva que cambie su inclinación en algún punto de negativa a positiva ó viceversa es no monótona.

El único modelo no monótono que ha sido usado con frecuencia supone que :

- 1.- El atributo es continuo y
- 2.- Cada reactivo tiene una curva que se aproxima a la distribución normal.

La probabilidad de responder en forma alfa es mayor en un punto X del atributo y ésta probabilidad de responder en la forma alfa se desploma en ambas direcciones y semeja la curva normal.

No es necesario que las curvas sean exactamente normales ni que sus desviaciones estándar sean idénticas. Este modelo ha sido empleado para un solo propósito, el desarro

lar ciertos tipos de escalas de actitudes.

El primer paso para elaborar una escala de ---- actitudes es contar con un gran número de afirmaciones sobre actitudes clasificadas por unos 100 jueces.

La esencia del modelo no monótono de Thurstone es que cada reactivo debe tender a recibir acuerdo unicamente en una zona del atributo. La mayor falla de la escala de ---- Thurstone y de otros modelos no monótonos de probabilidad es que es muy difícil encontrar reactivos que encajen en ellos, es obvio que el modelo no se ajusta a la mayor variedad de ---- reactivos. Está claro que este tipo de escala no se aplica -- a reactivos que requieren juicios.

Aún con respuestas referentes a sentimientos, - el modelo de Thurstone parece aplicable unicamente a ciertos tipos de exposiciones relacionadas con las actitudes y aún en esto el modelo tendría problemas lógicos. Solo las afirmaciones con dos contestaciones encajan aquí. Solo es posible construir tales reactivos mediante la sutil elaboración de 2 ó -- más declaraciones en lo que es ostensible que solo una es el reactivo. Este tipo de reactivos no solo es de difícil cons--- trucción sino que también tiende a ser ambiguo para los sujetos.

Otra crítica importante a los modelos no monó-- tonos de probabilidad es que es muy difícil pensar en los --- reactivos para los extremos de la escala. Resumiendo los modelos no monótonos de probabilidad se aplican supuestamente a - solo ciertos tipos de reactivos para la medición de las acti-

... y existen mejores formas para construir escalas de ---
actitudes. Los modelos monótonos con formas especificadas de
distribución; en algunos de estos modelos que se supone dan -
curvas monótonas, encajan dentro de una función normal. La --
característica más importante es que es más discriminatoria--
en ciertos niveles del atributo que lo es a niveles cercanos.
La buena zona de discriminación es por debajo de la parte que
asciende rápidamente a la curva. Mientras más acentuada es --
ésta parte mayor será la correlación biserial de ese reactivo
con el atributo. Si esta parte de la curva fuera vertical las
colas desaparecerían, el reactivo correlacionaría perfectamen
te con el atributo y con tales reactivos formaríamos una esca
la de Guttman.

Se ha encontrado que los juicios sobre los pe--
sos por el método del estímulo constante generalmente se ajust
ta a la curva normal.

La segunda razón para la atracción de este mode
lo es que tiene útiles propiedades matemáticas que permiten -
la deducción de muchos principios importantes. La suma de las
probabilidades de cualquier número de curvas normales sobre -
un punto, sobre un atributo sería la suma esperada de los re-
sultados de la prueba como conjunto, tomando en cuenta a las-
personas en ese punto. Esto significaría que la suma de los -
resultados sobre los reactivos para obtener resultados fina--
les, produce una escala que no ésta relacionada linealmente--
con el atributo. Cuando los reactivos se combinan y varían en
dificultad las curvas normales combinadas tienen menos forma-
de S que cuando todos los reactivos son iguales en dificultad.
Por estas dos razones aún aceptando el modelo de curva normal,
PAPEL PARA USO EXCLUSIVO DE S4 U. N. A. M.

es razonable suponer que los resultados totales de la prueba tienen una relación lineal aproximada con el atributo.

Otras deducciones interesantes de este modelo-- se refieren al monto del error en la medición - sin confiabilidad - en una prueba correspondiente a diferentes puntos sobre el atributo. La función que más uso a logrado es la curva logística. A simple vista la curva normal y la logística son similares, la ventaja de la curva logística es que es más fácil de trabajar matemáticamente.

Un punto importante que hay que tomar en cuenta en la discusión de modelos monótonos con formas especificadas de distribución, es que estos modelos no han conducido a escalar a personas en forma diferente al enfoque convencional --- que es la suma de los resultados sobre los reactivos. La situación es la misma con todos los demás atributos. Lo que estos modelos si permiten ciertas deducciones importante sumando resultados de los reactivos y han demostrado ser muy útiles para tales propósitos.

Finalmente llegamos al modelo monótono con formas de distribución no específica, que si responde a casi todos los esfuerzos por escalar personas y animales inferiores.

Hace tres suposiciones principales :

- 1.- Solo supone que cada reactivo tiene una curva monótonica.
- 2.- Se supone que las curvas de un conjunto particular de reactivos, en la curva del total de los resultados es aproximadamente lineal.

3.- La tercera suposición es que los reactivos, en conjunto, tienden a medir únicamente el atributo de que se trata. O sea que existe un factor común entre todos los reactivos.

Estas tres suposiciones constituyen el modelo lineal ó llamado de otra manera Modelo Sumativo, que es lineal por dos razones:

- 1.- Se supone que la suma de los resultados de los reactivos tienen una relación lineal -- aproximada con el atributo de que se trata.
- 2.- Es la más importante, el modelo lleva una combinación lineal de reactivos. Una sencilla suma de variables, es una combinación lineal de variables y una sencilla suma de resultados de reactivos, es una combinación lineal de estos resultados.

La forma más sensata de medir los atributos psicológicos de las personas es sumar los resultados sobre los reactivos. La esencia de lo lineal es que se toma muy en serio los reactivos individuales. Reconoce que el reactivo individual tiene especificidad considerable y errores de medición.

- B I B L I O G R A F I A
- 1.- Guilford, J.P. : "Psychometric Methods". Mc. Graw-Hill, 1954.
 - 2.- Guilford, J.P. : "Fundamental Statistics in ----- Psychology and Education". Mc. Graw Hill, 1965.
 - 3.- Gulliksen, H. and Messick, S. : "Psychological --- Scaling Theory and Applications". John Willey, ---- 1960.
 - 4.- Helmstadter, G.C. : "Principles of Psychological-- Measurement". Apleton Century Crofts, 1964.
 - 5.- Lord and Novick. : "Statistical Theories of Mental Test Scores". Addison - Wesley Publishing Cop., -- 1968.
 - 6.- Noll, H. Victor. : "Introduction to Educational -- Measurement". Haughton Mifflin Company, 1967.
 - 7.- Nunnally, C. Jun. : "Psychometric Theory". Mc. --- Graw Hill, 1967.
 - 8.- Siegel, Sidney. : "Nonparametric Statistics". Mc.- Graw Hill, 1956.
 - 9.- Thorndike, Robert L. y Hagen, Elizabeth. : "Test - y Técnicas de Medición en Psicología y Educación". Editorial Trillas, 1970.
 - 10.- Torgerson, S. Warren. : "Theory and Methods of ---

Scaling". John Wiley & Sons, 1967.

11.- Edwards, A. L. : "Techniques of Attitude Scale --
Construction". New York : Appleton Century Crofts,
1957. (a)