

Flem 8

leitura de approfondiments

pp. 160-230

TÉCNICAS DE INVESTIGACION  
APLICADAS A LAS  
CIENCIAS SOCIALES

JORGE PADUA  
INGVAR ARMAN  
HÉCTOR APÉCHEMA  
CARLOS BORSOTTI

EL COLEGIO DE MEXICO



FONDO DE CULTURA ECONOMICA  
México

Item 1: ¿Está Ud. entre los 170-174 cm de altura?	SI	(NO)
Item 2: ¿Está Ud. entre los 175-179 cm de altura?	SI	NO
Item 3: ¿Está Ud. entre los 180-184 cm de altura?	SI	(NO)
Item 4: ¿Está Ud. entre los 185-190 cm de altura?	SI	(NO)

Un ítem incluido en una serie de ítems diferenciados debe ser formulado de manera tal que sea respondido con un "SI" únicamente por las personas que tienen una posición fija a lo largo de la variable. Un ítem incluido en una serie de ítems acumulativos, debe ser formulado de manera tal que solamente pueda ser respondido con un "SI" a un lado de una determinada posición a lo largo de la variable investigada.

Trabajando con la altura de las personas, la diferencia entre estos dos tipos de serie de ítems aparece como trivial. Sin embargo, en el tratamiento de las escalas para medición de actitudes, su diferenciación es importante. Consistentemente, en cada tipo de las escalas a examinar (Lickert, Thurstone, Guttman, comparación por pares) tomaremos en cuenta la especificación de esta perspectiva.

#### ESCALAS PARA LA MEDICIÓN DE ACTITUDES

Vamos a utilizar la definición de escalas hecha por Stouffer.<sup>3</sup>

Se dice que existe escala cuando a partir de una distribución de frecuencias multivariada de un universo de atributos, es posible derivar una variable cuantitativa con la cual caracterizar los objetos de un modo tal que cada atributo sea una función simple de aquella variable cuantitativa.

Una escala es una forma particular de ítems, aunque aquí utilizaremos una serie de procedimientos objetivos para la selección de ítems de manera tal de controlar los errores producto de la subjetividad del investigador. (Ver el capítulo sobre "Conceptos, indicaciones e índices.") Construir una escala implica una serie de procedimientos mediante los cuales —de acuerdo a distintas reglas— se seleccionan ítems y se adjudican números a un conjunto de ítems (juicios o sentencias), número que va a expresar la intensidad que un sujeto o grupo de sujetos manifiestan en la variable.

Las actitudes en el contexto individual representan un estado mental que es un puente entre estados psicológicos y objetos exteriores. Kretsch y Cruschfield<sup>4</sup> sostienen a este respecto que se puede definir a una actitud "como

<sup>3</sup> Stouffer, S. et al.: *Studies in Social Psychology in World War II. Measurement and Prediction* (vol. IV); John Wiley & Sons, Nueva York, 1956.

<sup>4</sup> Kretsch D.; Cruschfield, R. S.: *Theory and Problems of Social Psychology*; McGraw-Hill, Nueva York, 1948.

una organización durable de procesos motivacionales, emocionales, perceptuales y cognitivos con respecto a algún aspecto del mundo del individuo". Las actitudes serían entonces procesos claves para entender las tendencias del individuo en relación con objetos y valores del mundo externo, aunque esas tendencias no son estáticas y, como las de Newcomb,<sup>5</sup> las actitudes representan un residuo de la experiencia anterior del sujeto. Las actitudes perdurarían en el sentido que tales residuos "son trasladados a nuevas situaciones, pero cambian en la medida en que nuevos residuos son adquiridos a través de experiencias en situaciones nuevas".

Las actitudes serían entonces tendencias a actuar con respecto a alguna entidad especificable (Newcomb); o como lo quieren Thomas y Znaniecki: "la tendencia individual a reaccionar, positiva o negativamente, a un valor social dado".

Las actitudes medidas por escalas deben interpretarse en términos analíticos no como "hechos", sino como "síntomas". Existe una serie de conceptos relacionados a las actitudes; entre ellos detallamos los siguientes:

*Creencia*: actitudes que incorporan una cantidad importante de estructura cognitiva. Las actitudes son *hacia* algo, mientras que las creencias son *en* o *sobre* algo.

*Sesgo (bias)*: son actitudes o prejuicios débiles, basados en premisas incompletas, deducidas falsamente o preconcebidas. Por lo tanto son poco precisas y relativamente fáciles de cambiar.

*Doctrina*: son objetos estímulos elaborados, hacia los cuales el individuo manifiesta una actitud. Una doctrina (republicana, liberal, católica, comunista, etc.) describe específicamente las razones para adherencia; por lo tanto una doctrina se aprende.

*Fe*: implica una actitud con alta carga emocional o afectiva. El sistema de actitudes referentes a la fe, describe una creencia fundamental y específica de la persona. La fe se ubica entre la creencia y la ideología.

*Ideología*: es un sistema cognitivo elaborado, que sirve para justificación de formas específicas de comportamiento, o como medio de racionalización. Ideología es concebido como un sistema lógico falso. La ideología se acepta como una fe.

*Valor*: en un sentido psicológico amplio, valores son marcos de referencia que sirven de guía o mapa para la evaluación de la experiencia y la conducta. Sistema de valores, sería la organización elaborada y articulada de actitudes, que toman valencias positivas, negativas o neutras en cuanto a objetos, estímulos, del ambiente y la relación de éstos con las metas de vida. *Opinión*: son evaluaciones tentativas, no fijas, sujetas a cambio o inversión. Es decir, son menos fijas y no comprometedoras para el individuo.

<sup>5</sup> Newcomb, T. M.: *Personality and Social Change: attitude formation in a student community*; Dryden, Nueva York, 1943.

<sup>6</sup> Thomas, W. I. y Znaniecki, F.: *The Polish peasant in Europe and America*.

ACTITUDES E INTERESES: ALCO ACERCA DE SU DIRECCIÓN E INTENSIDAD

Señalamos brevemente aquí algunas propiedades de las actitudes que tienen que ver con los problemas asociados a la medición de las mismas a través de escalas.

- 1) Dirección. La actitud puede ser positiva o negativa. En algunos casos es explicable hablar de una actitud neutral cuando el sujeto no es ni positivo ni negativo hacia el objeto.
- 2) Intensidad. La intensidad de la actitud es alta si el sujeto está fuertemente convencido que la actitud es justificada y baja si el sujeto no piensa así. Al respecto, una medida aceptable es la frecuencia con la cual el sujeto usa fuertes expresiones que señalan un engrane emocional y en la medida en que él está preparado a argumentar en favor de su posición.
- 3) Estabilidad. La actitud es estable si permanece invariable por un período muy largo.
- 4) Fortaleza. Una actitud es fuerte si, por ejemplo, es difícil de cambiar a través de persuasión o propaganda, y suave si cambia fácilmente.
- 5) Importancia. La importancia de la actitud es grande, si influye sobre la conducta de una persona en muchos campos de actividades.
- 6) Visibilidad, observabilidad o relevancia externa. Esta es alta si es fácil concluir a partir de observaciones sobre la conducta del sujeto (manera de hablar, acciones) que él tiene cierta actitud.
- 7) Relevancia interna. Es detectable si en la experiencia del sujeto la actitud por sí misma tiene una parte importante en su mundo de referencia. La actitud de una persona hacia la religión puede tener baja relevancia externa y, al mismo tiempo, una alta relevancia interna.
- 8) Involucramiento del ego. Cuando una actitud llega a constituir una parte importante de la personalidad, o es algo adherido a ella.
- 9) Integración y aislamiento. De un sistema mayor de actitudes interrelacionadas (perspectiva de vida, ideologías). Las actitudes políticas o religiosas están frecuentemente incluidas en tal sistema complejo, mientras que una actitud hacia una cierta pasta de dientes es experimentada como aislada, sin relación lógica o psicológica al complejo sistema total.
- 10) Especificidad o precisión. La imaginación es dirigida hacia la actitud. Ejemplo: un profesor de teología tiene una actitud más específica hacia la Iglesia católica que el "hombre de la calle".
- 11) Verificabilidad. La imaginación es igualmente dirigida hacia la actitud. El conocimiento es verificable, la fe, no; las opiniones ocupan un lugar intermedio y pueden a veces ser verificadas.

Debe notarse que muchas de las propiedades de las actitudes están correlacionadas. Alta intensidad frecuentemente significa estabilidad, fortaleza, alta relevancia interna y externa y alto grado de involucramiento del ego. Es importante observar los casos en los cuales las propiedades no están combinadas. En dichos casos, las distinciones hechas anteriormente son obvias. Las pruebas ordinarias de actitud sólo admitten la medición de algunas

de las propiedades mencionadas, principalmente dirección e intensidad. Para determinar las otras propiedades de la actitud podría ser útil una entrevista profunda del tipo entrevista focalizada.

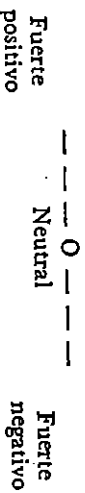
Dirección e intensidad frecuentemente son vistas como una sola propiedad de la actitud. Este enfoque es justificado por lo siguiente: suponiendo que dirección e intensidad son dos variables independientes, si de cada una de ellas pueden obtenerse tres valores diferentes, las combinaciones siguientes son posibles:

Cuadro 1

Intensidad	Dirección		Neutral	Negativo
	Positivo	Fuerte		
Fuerte	+++	+	000	---
Ni fuerte ni débil	+		00	--
Débil			0	

Una serie de investigaciones (entre otras la realizada por Edward H. Suchman) han mostrado, sin embargo, que una actitud neutral demuestra una baja intensidad. En la mayoría de los casos, las dos áreas rayadas podrían ser consideradas como inexistentes.

Los resultados son siete combinaciones, las cuales pueden ser ordenadas de la siguiente manera:



Esta perspectiva ha sido también usada en la construcción de un significativo "punto cero" para las escalas de actitudes. El Cuadro 2 resume la semejanza y diferencias entre tres de las principales escalas y el tipo de enfoque utilizado en ellas. En la parte final del capítulo agregamos el método de comparación por pares, de diferencial semántico y de escala de distancia social de Bogardus.

LA ESCALA LICKERT

Es un tipo de escala aditiva que corresponde a un nivel de medición ordinal, consistente en una serie de ítems o juicios ante los cuales se solicita la reacción del sujeto. El estímulo (ítem o sentencia) que se presenta al sujeto y las respuestas son solicitadas en términos de grados de acuerdo o desacuerdo que el sujeto tenga con la sentencia en particular.

CUADRO 2. *Semejanzas y diferencias entre algunos tipos de escalas para la medición de actitudes y en los enfoques Centrado en el sujeto, Centrado en el estímulo y Response-approach*

Semejanzas	DIFERENCIAS		
	Centrado en el sujeto	Centrado en el estímulo	Response-approach
Los tres tipos de enfoques se usan en la construcción de escalas que permiten medir distinciones de grado, más que de cualidad entre individuos.	Consiste en preguntarle a un sujeto su opinión acerca de un objeto para que él se ubique en determinado punto de la escala.	Consiste en preguntarle dónde ubicaría un estímulo sobre un continuo que representa el objeto de estudio.	El propósito puede ser escalar los sujetos, los estímulos, o ambos en relación a un objeto.
	Se dan valores a los sujetos.	Se dan valores a los estímulos.	Se dan valores tanto a los sujetos como a los estímulos.
	Se subrayan las diferencias individuales.	Se tratan de eliminar las diferencias individuales.	Se tienen en cuenta las diferencias en los sujetos y en los individuos.
	Análisis de la varianza unidireccional.	Análisis de la varianza unidireccional.	Análisis de la varianza en dos direcciones.
<i>Semejanzas en las escalas Lickert-Thurstone-Guttman.</i>	<i>Ejemplo: escala Lickert</i>	<i>Ejemplo: escala Thurstone</i>	<i>Ejemplo: escala Guttman</i>
Con pequeñas modificaciones los mismos <i>items</i> pueden utilizarse en los diferentes tipos de escala, ya que la	Escala ordinal.	Pretende ser intervalar.	Escala ordinaria.
	Escala aditiva.	Escala diferencial.	Escala acumulativa.
diferencia entre las escalas no reside en la elección de los <i>items</i> , sino en la relación lógica existente entre ellos.	Pretende ser unidimensional, pero de hecho se mezclan dimensiones.	Idem. a Lickert.	Es unidimensional, ésta es determinada por el coeficiente de reproductividad.
Las tres emplean el método de jueces en la construcción de la escala.	Los <i>items</i> se elaboran unidireccionalmente, procurando un 50 % de <i>items</i> positivos y un 50 % de <i>items</i> negativos.	Los <i>items</i> se formulan de manera que $\frac{1}{3}$ sean positivos; $\frac{1}{3}$ negativos y $\frac{1}{3}$ neutros.	Como es una técnica de análisis, los <i>items</i> se formulan de acuerdo a la técnica que se decida utilizar, Lickert o Thurstone.
Estrictamente todas las escalas miden a nivel ordinal.	Los <i>items</i> son acumulativos, por ello pueden estar aproximadamente en la misma posición de la escala.	Los <i>items</i> son diferenciadores.	Cada <i>item</i> tiene el carácter de acumulativo.
En principio las tres son unidimensionales.	La cantidad de <i>items</i> en la versión de los jueces varía de 30 a 50. En la versión final quedan entre 15 y 25.	El número de <i>items</i> iniciales puede llegar a los 200. En la escala final pueden quedar un máximo de 50.	La cantidad de <i>item</i> depende de la técnica de construcción.
En el momento de aplicar la escala ya validada (versión final), aunque se hayan utilizado distintos enfoques en la construcción, todos los procedimientos van a estar centrados en el sujeto. Lo que subyace en todos los procedimientos es el obtener conocimientos de las actitudes que forman parte de un contexto sociocultural.	La selección de los <i>items</i> se realiza en base a su poder discriminativo.	La selección de los <i>items</i> se hace en base al recorrido intercuartílico.	La selección de los <i>items</i> se basa en su escalabilidad, que se determina en base al número de errores.
	Las alternativas de respuesta en cada <i>item</i> puede variar de 3, 4, 5 o más alternativas.	El puntaje de los <i>items</i> se determina por el valor de la mediana de los puntajes asignados por jueces, sobre una escala que va de 1 a 11.	El puntaje de los <i>items</i> varía de acuerdo a la técnica empleada.
	No existe gradación de los <i>items</i> a lo largo de un continuo.	Se presentan mezclados <i>items</i> positivos, negativos y neutros.	Los <i>items</i> se ordenan en forma decreciente de acuerdo a su grado de dificultad.

Semejanzas	DIFERENCIAS		
	Centrado en el sujeto	Centrado en el estímulo	Response-approach
El número de jueces varía de 50 a 100. Éstos deben tener características similares a las de los sujetos en el universo a estudiar.	El número de jueces es entre 50 y 200. Se exige de ellos objetividad e información.	El número de jueces depende de la técnica empleada en la construcción.	
La consistencia interna de la escala se establece mediante el método de correlación por mitades ( <i>split-half</i> ).	La consistencia interna se basa en el cálculo de la correlación entre cada <i>item</i> con el puntaje total del <i>test</i> ( <i>item-test</i> ).	La consistencia interna está determinada por la escalabilidad, calculada en base a la diferencia $CR-MMR$ .	
A cantidad igual de <i>items</i> es más confiable que la escala Thurstone. La confiabilidad aumenta con el incremento en las alternativas de respuesta.	Cuando se alcanza un número de 50 <i>items</i> es más confiable que la Lickert.	Es más confiable debido a su unidimensionalidad.	
Es más fácil y rápida de construir.	Difícil de construir. Gasto fuerte en términos de tiempo y trabajo.	Depende de la técnica utilizada en la construcción.	
Escala correspondiente a <i>test</i> de Tipo A.	Corresponde a un <i>test</i> de Tipo B.	Corresponde a ambos tipos de <i>test</i> (A y B).	
En la escala final se presentan los <i>items</i> con la cantidad de alternativa idéntica a la de la versión de los jueces.	En la escala final se presentan solamente dos alternativas de respuesta: acuerdo-desacuerdo.	En la escala final se presentan los <i>items</i> en orden de dificultad creciente.	
El puntaje máximo es igual al número de <i>items</i> multiplicado por el puntaje mayor en cada alternativa de respuesta: el puntaje mínimo es igual al número de <i>items</i> multiplicado por el puntaje menor en las alternativas de respuesta. Para la ubicación de individuos se pueden utilizar también en valores promedio.	El puntaje máximo y mínimo en la escala se determinan por la sumatoria de <i>items</i> ponderados.	Los sujetos se ubican en la escala en forma decreciente: desde los que contestaron en forma positiva a todas las preguntas, hasta los que contestaron en forma negativa a todas las preguntas.	
Con frecuencia la puntuación total de un individuo puede tener un significado poco claro, cuando se lo compara con otros individuos, ya que combinaciones distintas pueden producir el mismo resultado.	Idem. a la Lickert.	Los individuos que tienen un puntaje de actitud más favorable, deben también tener una actitud más favorable en cada <i>item</i> escalado.	
El problema principal de la escala es la validez. Determinar cuándo una misma puntuación alcanzada por combinación de distintas categorías de respuesta tiene consecuencias para la interpretación del atributo en cuestión y cuándo no.	El problema principal es el centro de la escala. Éste nos dice mucho acerca del significado que tiene el hecho que un individuo ocupe una posición en el centro de la escala.	En el caso de actitudes complejas, no es muy eficaz.	

El método del *summated ratings* de Lickert resulta de la suma algebraica de las respuestas de individuos a ítems seleccionados, previamente como válidos y confiables. Si bien la escala es aditiva, no se trata de encontrar ítems que se distribuyan uniformemente sobre un continuo "favorable-desfavorable", sino que el método de selección y construcción de la escala apunta a la utilización de ítems que son definitivamente favorables o desfavorables con relación al objeto de estudio. El puntaje final del sujeto es interpretado como su posición en una escala de actitudes que expresa un continuo con respecto al objeto de estudio.

La escala Lickert es pues un *test de Tipo A*, ya que el sujeto, a través de su conducta en la situación de prueba, conscientemente o inconscientemente, se sitúa a lo largo de la variable. La escala Lickert es también una escala del tipo *centrada en el sujeto (subject-centered)*: el supuesto subyacente es que la variación en las respuestas será debida a diferencias individuales en los sujetos. Veremos además cuando examinemos los pasos en la construcción de la escala, que la escala inicial se administra a una muestra de sujetos que actuarán como *juces*. (Esa muestra de sujetos debe ser representativa de la población a la que se aplicará la escala final.)

Finalmente, los ítems son seleccionados en base a su *poder discriminatorio* entre grupos con valores altos y con valores bajos en la variable. Es decir, que lo que interesa es la coherencia, entendida ésta en función de las respuestas.

#### A) La construcción de una escala Lickert

La construcción de una escala de este tipo implica los siguientes pasos: 1) Es necesario construir una serie de ítems relevantes a la actitud que se quiere medir. 2) Los ítems deben ser administrados a una muestra de sujetos que van a actuar como *juces*. 3) Se asignan puntuajes a los ítems según la dirección positiva o negativa del ítem. 4) Se asignan los puntuajes totales a los sujetos de acuerdo al tipo de respuesta en cada ítem, la suma es algebraica. 5) Se efectúa un análisis de ítems. 6) Se construye con base en los ítems seleccionados la escala final. Examinemos en detalle cada uno de los pasos:

##### 1) La construcción de los ítems

Los ítems que van a aplicarse a la muestra inicial de jueces, cuyo número debe ser entre 30 y 50. Para la construcción de los ítems deben tomarse en cuenta los siguientes criterios, que aparecen en Edwards:<sup>7</sup>

- a) Evite los ítems que apuntan al pasado o al presente.
- b) Evite los ítems que dan demasiada información sobre hechos, o aquellos que pueden ser interpretados como tales.
- c) Evite los ítems ambiguos.
- d) Evite los ítems irrelevantes con respecto a la actitud que quiere medir.
- e) Los ítems en la escala deben formularse según expresen actitudes o ju-

<sup>7</sup> Edwards, A. L.: *Techniques of Attitude Scale Construction*, Appleton-Century-Crofts: Nueva York, 1957.

cios favorables o desfavorables con respecto a la actitud. No se trata de elegir ítems que expresen distintos puntos en el continuo.

- f) Evite los ítems con los cuales todos o prácticamente nadie concuerda.
- g) Los ítems deben ser formulados en lenguaje simple, claro y directo.
- h) Solamente en casos excepcionales exceda de las 20 palabras cuando formule el ítem.
- i) Un ítem debe contener sólo una frase lógica.
- j) Los ítems que incluyan palabras como "todos", "siempre", "nadie", etc. deben omitirse.
- k) De ser posible, los ítems deben ser formulados con frases simples, y no compuestas.

l) Use palabras que el entrevistado pueda comprender.

m) Evite las negaciones, particularmente las dobles negaciones.

n) Combine los ítems formulados positiva y negativamente en una proporción aproximada a 50% — 50%.

Un sistema que puede ser aplicado para eliminar muchos ítems dudosos o que dan demasiados hechos es el siguiente: cada miembro del grupo de investigación responde a los ítems asumiendo primero una actitud positiva hacia la variable y luego responde como si tuviese una actitud negativa. Si la respuesta en ambos casos se ubica en la misma categoría, el ítem no es apropiado para incluirse en la versión de los jueces.

Cada ítem es entonces un juicio o una sentencia a la cual el juez debe expresar su grado de acuerdo o desacuerdo. La graduación de acuerdos o desacuerdos varía en cantidad de alternativas que se le ofrece al sujeto; éstas pueden ser 3, 4, 5, 6 o 7 alternativas. En general la decisión sobre la cantidad de alternativas a ofrecer dependerá no tanto de las "preferencias personales" del investigador, sino del tipo de investigación, del tipo de pregunta, del tipo de distribución de la variable, etc. Ejemplificaremos a continuación algunos ítems con sus respectivas alternativas.

Las siguientes afirmaciones son opiniones con respecto a las cuales algunas personas están de acuerdo y otras en desacuerdo. Indique, por favor (marcando con una X en el paréntesis correspondiente), la alternativa que más se asemeja a su opinión.

- 1) Las mujeres no deberían meterse en política
 

( )	Muy de acuerdo
( )	De acuerdo
( )	Ni acuerdo ni desacuerdo
( )	En desacuerdo
( )	Muy en desacuerdo
- 2) Leyendo lo que se publica en los diarios y las informaciones de radio y TV, es posible tener una idea acertada de lo que ocurre en la situación política mexicana
 

( )	Muy de acuerdo
( )	De acuerdo
( )	En desacuerdo
( )	Muy en desacuerdo
( )	Ni acuerdo ni desacuerdo

- 3) Los manifiestos, proclamas y solicitudes que publican en los diarios los partidos políticos no informan sobre sus verdaderos propósitos
- ( ) Totalmente de acuerdo  
 ( ) Medianamente de acuerdo  
 ( ) Escasamente de acuerdo  
 ( ) Medianamente en desacuerdo  
 ( ) Totalmente en desacuerdo

- 4) Vivo intencionalmente el presente sin pensar en el futuro.
- ( ) Verdadero  
 ( ) Falso

Los ejemplos expresan *ítems* positivos y negativos; manifiestos y latentes y con distintas alternativas de respuesta. Los *ítems* son construidos a partir de juicios que expresan alguna relación postulada a nivel de la teoría sustantiva, y de observaciones empíricas de afirmaciones de grupos o sujetos que pertenecen a grupos o asociaciones que manifiestan la propiedad que se quiere medir. Los *ítems* así pueden ser extraídos de libros, publicaciones y artículos que tratan teóricamente sobre el objeto que se quiere medir. También puede el investigador acudir a análisis de contenidos sobre discursos o manifiestos de individuos y asociaciones (por ejemplo si se trata de medir radicalismo-conservadurismo, una fuente muy rica para la formulación de *ítems* son las declaraciones de grupos de interés: empresarios, grupos políticos de izquierda y de derecha, etc.). Otras estrategias para la construcción de *ítems* aparecen señaladas en la sección correspondiente a la escala Thurstone.

#### 2) La administración de los ítems a una muestra de jueces

Una vez construidos los *ítems* (30 a 50) éstos van a ser distribuidos a una muestra de jueces (entre 50 a 100) los cuales deben ser seleccionados al azar de una población con características similares a aquella en la cual queremos aplicar la escala final. (Para los procedimientos de selección de la muestra ver el capítulo *Muestra* en este manual.) Estos jueces responderán a cada uno de estos *ítems* su opinión. Las instrucciones a los jueces pueden ser dadas según el siguiente ejemplo:

##### EJEMPLO DE UNA VERSIÓN PRELIMINAR

El presente es un estudio de opiniones de estudiantes universitarios respecto a problemas actuales de la universidad.

A continuación se le presentará una serie de afirmaciones respecto a las cuales algunas personas están de acuerdo y otras en desacuerdo. Después de cada afirmación se presentarán cinco alternativas de respuestas posibles:

- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo.

Indique por favor —marcando con una cruz entre el paréntesis— la alternativa que más se asemeje a su opinión. Cuando no entienda alguna afirmación, ponga un signo de interrogación (?) al frente de ella. Trate de responder lo más rápido posible. Muchas gracias.

1. Las representaciones estudiantiles deberían participar en las decisiones sobre planes de estudio.
- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

2. Las clases en las que el profesor tiene todo el control son las que mejor resultados producen en el aprendizaje.
- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

3. El plan de estudios de la UNAM debe ser centralizado en la Secretaría de Educación Pública.
- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

4. El trabajo en grupo es más productivo que el trabajo individual.
- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

5. Las carreras a las que el gobierno y la Universidad deberían prestarles más apoyo son aquellas contradas en las necesidades del país.
- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

6. El alumno debe tener libertad en la elección de cuál es la mejor manera de controlar su rendimiento académico.
- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

7. La única obligación de los alumnos es estudiar. Los planes de estudio son asunto de los profesores.
- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

8. La mejor manera de juzgar a un estudiante es por su rendimiento académico.
- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo

8. En la asignatura de Física, los estudiantes que se dedican a estudiar en la universidad deberían tener un nivel de conocimientos más alto que los que se dedican a estudiar en la escuela secundaria. ( ) Totalmente en desacuerdo ( ) Totalmente de acuerdo
9. Un trabajo hecho con consulta (en equipo o individualmente) permite una mejor evaluación de los conocimientos de los alumnos que una prueba hecha en la clase. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
10. Es preferible que cada escuela universitaria tenga un programa fijo de estudios en vez de que, como sucede en otros países, el alumno pueda escoger con alguna libertad ciertas materias de su agrado. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
11. Es preferible que los alumnos no hagan preguntas o intervenciones durante la exposición del profesor. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
12. No conviene que los alumnos intervengan en la confección de los programas de estudio. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
13. Las pruebas y exámenes deben limitarse exclusivamente a evaluar el grado de conocimiento de los alumnos respecto a la materia expuesta durante las horas de clase. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
14. Es inconveniente que ocupen plazas en la universidad estudiantes que se verán impedidos de seguir sus estudios por falta de medios económicos. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
15. Deben tener acceso a la cátedra universitaria sólo personalidades científicas de determinada orientación ideológica. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
16. No habría por qué hacer esfuerzos en adecuar los horarios de clase para la gente que trabaja. O se trabaja o se estudia. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo

17. No deberían otorgarse becas a alumnos capaces, pero con recursos económicos suficientes, sino a aquellos de escasos recursos, aunque sean menos capaces. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
18. Sólo a los organismos centrales de dirección de la enseñanza universitaria les está dada la facultad de decidir a que ramos de la enseñanza han de conceder becas de estudios. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
19. En el organismo universitario destinado a la distribución de becas de estudio entre las distintas facultades no deben participar estudiantes. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
20. Para un mejor aprovechamiento de los recursos debe haber un organismo central que controle la concesión de becas de estudios. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
21. A la hora de tomar un acuerdo importante sobre las evaluaciones docentes universitarias, las autoridades deben hacerlo sin tener en cuenta la opinión de los alumnos. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
22. La universidad debería exigir de los egresados una retribución por los estudios recibidos, estableciendo un impuesto por el ejercicio profesional. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
23. No compete a las escuelas universitarias fijar la cantidad de matriculas anuales para sus alumnos, sino a un organismo superior central. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
24. La decisión respecto a la selección y otorgamiento a los estudiantes de material académico, debe ser tomada exclusivamente por los organismos centrales de la universidad. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo
25. La ampliación de las carreras universitarias existentes, debe responder solamente a las necesidades de la sociedad. ( ) Totalmente de acuerdo ( ) De acuerdo en general ( ) Ni de acuerdo ni en desacuerdo ( ) En desacuerdo en general ( ) Totalmente en desacuerdo



lamente al número de los postulantes que se presenten en cada escuela.

- ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

26. En las pruebas y exámenes escritos no debe haber personal universitario que vigile a los alumnos.

- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

27. Los servicios de habitación y restaurante brindados por la universidad deben ser administrados sin ninguna participación del estudiantado.

- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

28. No debería haber un organismo central de planificación universitaria.

- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

29. Los alumnos no deben tener injerencia alguna en la labor del personal docente.

- ( ) Totalmente de acuerdo  
 ( ) De acuerdo en general  
 ( ) Ni de acuerdo ni en desacuerdo  
 ( ) En desacuerdo en general  
 ( ) Totalmente en desacuerdo

### 8) Asignación de puntajes a los ítems

Con este paso comienza efectivamente el análisis de la escala. Hay que clasificar a cada ítem según sea positivo o negativo, y luego ponderar las alternativas de respuesta. Nuevamente existen diferentes criterios para la adjudicación de las ponderaciones. Por ejemplo, los pesos para un ítem positivo pueden ser:

- Pesos
- 4 ( ) Totalmente de acuerdo  
 3 ( ) De acuerdo en general  
 2 ( ) Ni de acuerdo ni en desacuerdo  
 1 ( ) En desacuerdo en general  
 0 ( ) Totalmente en desacuerdo

o la alternativa:

- Pesos
- 2 ( ) Totalmente de acuerdo  
 1 ( ) De acuerdo en general

- 0 ( ) Ni de acuerdo ni en desacuerdo  
 -1 ( ) En desacuerdo en general  
 -2 ( ) Totalmente en desacuerdo

o cualquier otra serie de números.

Por lo general desaconsejamos la utilización de signos positivos y negativos en la adjudicación de los puntajes o de pesos a las alternativas de respuestas, ya que pueden crear la falsa impresión que la escala está midiendo a nivel intervalar; esto es, donde tendríamos puntajes finales en los que existe una posición 0, posiciones +1, +2... +40; y posiciones -1, -2, -3... -40. De hecho la escala mide a nivel ordinal y los valores de escala, simplemente implican posiciones de rango.

Para los ítems negativos, hay que recordar que la serie de números a adjudicar debe ser inversa. Por ejemplo, en un ítem negativo, la ponderación siguiendo la primera alternativa se haría:

- Pesos
- 0 ( ) Totalmente de acuerdo  
 1 ( ) De acuerdo en general  
 2 ( ) Ni de acuerdo ni en desacuerdo  
 3 ( ) En desacuerdo en general  
 4 ( ) Totalmente en desacuerdo

Los ítems se ubican ya sea en forma positiva o negativa en relación a la variable con el fin de controlar los efectos del "response-set", esto es, controlar las pausas de respuesta de aquellos respondientes que tienden a dar respuestas afirmativas o negativas de manera automática.

### 4) Asignación de puntajes totales

Este paso consiste simplemente en la adjudicación de los puntajes totales para cada individuo en la muestra de jueces. Esta suma resultará de la adición de los puntajes ponderados para cada ítem. En el caso de ítems con valores negativos, la suma es algebraica.

De comienzo estamos asumiendo que las personas con alto grado en la variable van a tener puntajes altos, mientras que las personas con una baja actitud manifestarán puntajes bajos. Si hemos presentado 30 ítems con un valor ponderado máximo de 4, y un mínimo de cero, la amplitud total de la dispersión de la variable a esperar sería entonces 120 (puntajes máximos de 120 y mínimo de 0 respectivamente).

### 5) Análisis de los ítems

Una vez computados los puntajes totales para todos los jueces, hay que ordenarlos de manera que el sujeto con el puntaje total más alto ocupe el

primer lugar, el segundo puntaje más elevado a continuación, etc., hasta llegar a la persona con el puntaje más bajo.

Una vez ordenados los sujetos, vamos a operar únicamente con los cuartiles superiores e inferiores es decir el 25% de los sujetos con puntajes más elevados y el 25% de los sujetos con puntajes más bajos. Del 50% del centro no nos vamos a preocupar más. Formamos de esta manera un grupo alto y un grupo bajo con respecto a la variable y a los puntajes totales. Si tuvimos 52 jueces, el grupo alto estará constituido por los 13 jueces con los puntajes más elevados y los 13 con puntajes más bajos.

Tomamos a estos 26 sujetos y los colocamos en una tabla en donde situamos las puntuaciones en cada ítem y el puntaje total para cada uno de los sujetos ordenados.

Hay que seleccionar ahora los ítems que discriminen mejor. Hay tres técnicas más en uso para la selección de los ítems: la del cálculo del poder discriminativo de cada ítem; la de correlación ítem-test; y el test de la mediana.

El procedimiento para el cálculo del poder discriminativo de un ítem, sigue la forma mencionada en el cuadro 3. Una vez separados el grupo alto y el grupo bajo, se calculan los promedios de cada ítem en cada uno de los grupos. Siguiendo el ejemplo de la figura 1, el promedio del ítem 1 es de 3.7 para el grupo alto (resultado de 48/13); y de 0.9 para el grupo bajo (resultado de 12/13).<sup>8</sup>

Una vez calculados los valores promedios para cada ítem en los grupos alto y bajo, procedemos a calcular el poder discriminativo de cada ítem según la fórmula:

$$DM = \frac{M_1 - M_2}{N_1 - 1 + N_2 - 1}$$

Donde:

t = Test t de Student

DM = Diferencia entre medidas (M<sub>1</sub> - M<sub>2</sub>)

s<sup>2</sup> = variancias de muestra 1 y 2, respectivamente

N = cantidad de casos en cada una de las muestras

Para el cálculo del poder discriminativo del ítem conviene utilizar el siguiente cuadro, que es una continuación del cuadro 3:

Ordenamiento de jueces	Ítem núm.																														Puntaje total	
	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	22	24	26	28	30												
1	4	4	4	4	4	1	4	2	4	3	4	4	4	4	etc...																	112
2	4					2		2																								108
3	3					4		0																								108
4	4					0		1																								104
5	4					1		2																								102
6	4					0		0																								101
7	4					0		0																								100
8	3					4		2																								98
9	4					2		1																								98
10	3					1		1																								98
11	3					3		1																								97
12	4					2		0																								97
13	4					2		2																								96
14	2					2		3																								27
15	0					3		4																								26
16	2					0		2																								26
17	1					1		4																								26
18	0					4		4																								24
19	1					0		3																								17
20	0					1		3																								16
21	0					1		3																								16
22	1					2		4																								14
23	2					1		4																								14
24	3					1		3																								14
25	0					2		4																								8
26	0					1		4																								7
27	0					4		3																								7

<sup>8</sup> Para llegar al promedio sumamos la columna correspondiente al ítem 1 en los 13 sujetos con puntajes altos dividiéndolos por el total de casos: (4+4+3+4+4+4+4+4+3+4+3+3+4+4)/13. En forma idéntica, para el grupo bajo el cálculo es: (2+0+2+1+0+1+0+1+2+3+0+0+0)/13.

Cuadro 4. Valores promedios para los grupos alto y bajo y diferencia de medias, para el cálculo del poder discriminatorio de ítems en la versión de los jueces en una escala Likert

	1	2	3	4	5	6	7	8	9	10	11	12
Promedio en el grupo alto ( $M_1$ )					1.8		1.1					
Promedio en el grupo bajo ( $M_2$ )					1.7		3.4					
Diferencia de medias ( $M_1 - M_2$ )					0.1		-2.3					

Una vez calculados los valores  $t$ , confrontamos con la tabla de distribución valores correspondientes, seleccionando aquellos ítems que realmente presenten diferencias significativas entre ambos grupos de contrastes.

En los ejemplos señalados en la figura 1, la diferencia para el ítem 1 es significativa al nivel de 0.1 y la del ítem 7 al nivel de 0.5. El ítem 5 debe ser desechado porque no discrimina significativamente. Es importante notar asimismo, que en el caso del ítem 7 hemos colocado el signo del ítem originalmente mal ya que, como vemos, el grupo con valores bajos responde al estímulo en forma más positiva que los del grupo alto. Corresponde entonces cambiar los valores, manteniendo al ítem.

El método *item-test* para la selección de los ítems consiste en correlacionar el puntaje del ítem con el puntaje total del *test*. El coeficiente de correlación a utilizar es el coeficiente de correlación biserial, ya que aquí se trata de dos variables, una que podemos considerar intervalar, y a la otra una serie dicotomizada, siendo además la dicotomía forzada. El supuesto que tenemos que aceptar es que la distribución original es continua y normal. La fórmula para el cálculo de la correlación será:

$$r_b = \frac{M_p - M_q}{\sigma_t} \cdot \frac{p \cdot q}{y}$$

Donde:

$M_p$  y  $M_q$  = Medias parciales en el grupo alto y bajo, respectivamente.

$\sigma_t$  = Desviación estándar total.

$p$  y  $q$  = Proporción de casos en una y otra distribución.

$y$  = Ordenada a la curva normal.

Cualquiera que sea la técnica utilizada para el análisis de los ítems, el objetivo es seleccionar aquellos que discriminen mejor (valores significativos de  $t$ , o valores altos de  $r_b$ ).

De hecho hay una alternativa al uso de la prueba  $t$ , o del coeficiente de correlación biserial, que estrictamente se corresponde más con el tipo de medición de nivel ordinal, con el que opera la escala Likert. En este caso la diferencia entre medianas, computada a través del "test de la mediana". Para ello hay que determinar primero el valor de la mediana de cada ítem para los grupos alto y bajo combinados, luego dicotomizamos los valores en una tabla de  $2 \times 2$  para cada ítem, de la siguiente forma:

	Grupo alto	Grupo bajo	
Número de puntajes por debajo de la mediana combinada	A	B	A + B
Número de puntajes por encima de la mediana combinada	C	D	C + D
	A + C	B + D	

A esta tabla aplicamos ya sea  $\chi^2$  o el test de Fisher, según sea la cantidad de casos (por lo menos  $N = 40$  para aplicar  $\chi^2$ ; con menos casos se recomienda el test de Fisher).

La fórmula para el cómputo de  $\chi^2$  para este tipo de figura es:

$$\chi^2 = \frac{N \left[ (AD - BC) - \frac{N}{2} \right]^2}{2}$$

Seleccionamos por supuesto aquellos ítems cuyo  $\chi^2$  da diferencias significativas.

Usando la misma figura, el test de Fisher se computa según la siguiente fórmula:

$$p = \frac{(A+B)! (C+D)! (A+C)! (B+D)!}{N! A! B! C! D!}$$

#### 6) La versión final de la escala

La cantidad de ítems seleccionados de acuerdo a su poder discriminativo, construirá la escala final a ser aplicada a sujetos o grupos como versión final. Los puntajes finales a adjudicar a los individuos, serán entonces el producto de la suma de los puntajes obtenidos en cada ítem, divididos entre el total de ítems.

#### B) Comentarios finales

Para calcular la confiabilidad de la escala, se puede utilizar la correlación entre mitades del test (*split-half reliability*): se correlacionan la suma de los

puntuajes en los ítems impares con la suma de los puntuajes de los ítems pares. Utilizamos  $p$  de Spearman, y luego la fórmula:

$$C = \frac{2p}{1+p}$$

A continuación presentamos las ventajas y desventajas de la escala Lickert, comparada con la Thurstone.

*Ventajas.* a) Permite la utilización de ítems que no se encuentran relacionados en forma manifiesta con la actitud que se desea estudiar (es decir, se pueden utilizar ítems con contenido latente). b) Es más rápida y fácil de construir. c) A mismo número de ítems, es más confiable. d) La cantidad de alternativas de respuesta permite una información más precisa de un sujeto en un ítem particular.

*Desventajas.* a) Por tratarse de una escala ordinal, no permite apreciar la distancia que hay entre pares de sujetos con respecto a la actitud. b) Con frecuencia dos puntuajes iguales pueden ocultar pausas de respuestas diferentes de los individuos. c) No hay garantía de unidimensionalidad; consecuentemente pueden mezclarse distintas dimensiones, no estando seguro el investigador de cuál de ellas realmente se trata.

#### LA ESCALA THURSTONE

Thurstone es quien provee la racionalidad —mediante su ley de juicios comparativos— para todo el aparato conceptual en la construcción de escalas para medir actitudes. Esta ley sostiene que para cada estímulo ( $e$ ) dado, está asociado un proceso modal discrimininal sobre un continuo psicológico. La distribución de todos estos procesos discriminacionales sigue la forma de la distribución normal, en la que todos los procesos discriminacionales producidos por el estímulo se distribuyen normalmente alrededor del proceso de discriminación modal, con una dispersión discrimininal ( $\sigma_d$ ). Dado un conjunto  $n$  de estímulos, es posible ordenarlos en un continuo psicológico tomando como referencia el grado de atributo que ellos poseen.

A partir de estos principios, Thurstone propone su *escala de intervalos aparentemente iguales*, de tipo diferencial, en la que los ítems son seleccionados por una serie de técnicas que permiten escalonarlos de manera tal que expresen el continuo psicológico subyacente. La medición trata de establecerse al nivel intervalar. Es decir, una escala en la que sea posible afirmar que la distancia que separa a un sujeto que obtuvo una puntuación de 8.7 con respecto a otro sujeto que obtuvo 6.3, es igual a la distancia que separa a otro par de sujetos que obtuvieron puntuaciones de 3.6 y 1.2 respectivamente, y a cualquier distancia que sea igual a 2.4 puntos. (Sin embargo, como veremos más adelante, es discutido que la escala mida efectivamente a este nivel.)

El continuo psicológico en la escala de intervalos aparentemente iguales de Thurstone, se edifica sobre una serie de juicios de actitud distribuidos

en una escala de 11 puntos, en la que el punto 1 de la escala representa una actitud extrema (favorable o desfavorable), el punto 6 representa una actitud neutra (ni favorable ni desfavorable); y el punto 11 el otro extremo (favorable o desfavorable, según el extremo contrario a la actitud asumida en 1).

Los ítems en la escala Thurstone son contruados, diseñados y seleccionados de manera tal que permitan atribuir a los sujetos a los que se aplicará definitivamente la escala, un punto en un continuo. Así, esta escala es un poco más refinada que la escala Lickert, e implica una cantidad considerable de trabajo adicional.

#### A) La construcción de una escala Thurstone

Los procedimientos para la construcción de una escala de este tipo son: 1) Se construye una serie de ítems (alrededor de 150). 2) Se solicita a un grupo de jueces (más o menos 100), que ubiquen a los ítems en una escala de 11 puntos. 3) Una vez evaluados los ítems por los jueces, se adjudica a los ítems valores de escala. 4) Se seleccionan los ítems que representan el rango entero de la escala, rechazando los ítems ambiguos.

Detallamos cada uno de los pasos:

#### 1) La construcción de los ítems (versión de los jueces)

Construya entre 100 y 200 ítems tomando en consideración los siguientes criterios (ver Edwards, A. L. para mayores detalles):

- Evite los ítems que señalan al pasado en vez del presente.
- Evite los ítems que dan demasiada información sobre hechos o los que fácilmente puedan ser interpretados como tales.
- Evite los ítems ambiguos.
- Evite los ítems irrelevantes con respecto a las actitudes que pretende medir.

e) Evite los ítems con los cuales todos o nadie concuerda.

f) Los ítems deben ser formulados en un lenguaje simple, claro y directo.

g) Sólo en casos excepcionales sobrepase las 20 palabras en un ítem.

h) Un ítem sólo debe contener una frase lógica.

i) Los ítems que incluyan palabras como "todos", "siempre", "nadie", "nunca", etc. serán perichidos de la misma manera y por ello deben omitirse.

j) De ser posible, los ítems deben ser formulados como frases simples y no compuestas.

k) Use sólo palabras que el entrevistado pueda comprender.

l) Evite las negaciones, especialmente las dobles negaciones.

m) Combine los ítems formulados positiva, neutral y negativamente en una proporción de 1/3, 1/3 y 1/3, distribuidos uniformemente sobre la variable.

Algunas maneras de formular ítems pueden ser: i) Extraerlos de libros, publicaciones y artículos que tratan sobre el objeto cuya actitud se quiere medir. Son importantes también las declaraciones y los discursos, a partir de los cuales uno pueda hacer una especie de análisis de contenido. Cuando se

sigue esta estrategia hay que tener en cuenta que un buen monto de reformulación va a ser necesario. *ii)* Concretar con respecto a la actitud. En este caso la presentación distintos puntos de vista con respecto a la actitud. En este caso la grabación de la discusión facilitará la selección de frases adecuadas. Aquí también la reformulación será necesaria. *iii)* Formular uno mismo o en cooperación con otros investigadores, los enunciados ante los cuales se espera que la gente reaccionará en forma positiva, negativa o neutra.

En todo caso nunca es fácil llegar a la enunciación de 100-200 ítems sin incurrir en repeticiones o en formulaciones muy similares. Es importante, por cierto, que la distribución de los ítems a ser presentados a los jueces sea aproximadamente pareja, conteniendo un tercio de ítems positivos, negativos y neutros, a lo largo de un continuo.

2) La administración a los jueces

La lista de ítems (100 a 200), se distribuye a jueces (preferiblemente 200, con un mínimo de 50), los cuales van a ubicar a los ítems en una escala intervalo-subjetiva que va de 1 a 11 puntos.

Los jueces son seleccionados en función de su conocimiento sobre el problema que se quiera medir; y en la clasificación o evaluación de los ítems no importa la opinión personal del juez, sino su evaluación del punto en la escala continua de 1 a 11 en el cual él ubica al ítem (es decir, la determinación del peso que el ítem tiene en su opinión para la medición de la actitud).

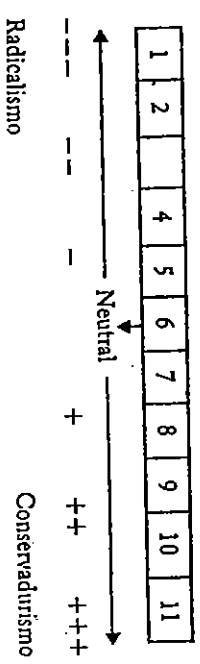
Las instrucciones a los jueces —que operan en forma independiente— pueden ser las siguientes:

EJEMPLO: ESCALA ACTITUD TIPO THURSTONE

Instrucciones

El Seminario de Recolección y Análisis de Datos del Centro de Estudios Sociológicos de El Colegio de México está realizando una serie de ejercicios sobre construcción de escalas para la medición de actitudes.

En el presente caso, tiene usted en sus manos una serie de afirmaciones para construir una escala de tipo Thurstone. La escala de intervalos aparentemente iguales de Thurstone parte de una serie de supuestos y técnicas en la cual los ítems son escalonados de manera tal que expresen un continuo subyacente. El continuo se edifica sobre una serie de juicios de actitud distribuidos en una escala de 11 puntos.



El punto 1 representa una actitud extrema (en nuestro caso: radicalismo), el punto 6 una actitud neutra (ni radical ni conservador) y el punto 11 el extremo conservador.

Para la construcción de esta escala, se requiere someter una cantidad de ítems a un número limitado de jueces antes de ser aplicada a una muestra de una determinada población.

Los jueces se eligen entre personas que tienen conocimientos especializados sobre la variable que se trata de medir, como ha sido el caso con usted.

En el presente caso, se trata de construir una escala que mida actitudes de radicalismo-conservadurismo. Quisieramos pedirle que, en relación a esto, no vuelque usted sus propias opiniones acerca de las afirmaciones que aparecen a continuación, sino que usted exprese su juicio acerca de cuán radical, conservadora o neutra le parece cada una de las afirmaciones.

A la izquierda de cada afirmación hay un cuadrado en el cual usted debe colocar, de acuerdo a su criterio, el número que representara la frase en el continuo de 1 a 11, siendo:

- 1 el extremo "radical",
- 6 el punto medio o "neutral",
- 11 el extremo "conservador".

Si cree que la expresión se ubica entre alguno de estos puntos, utilice el número intermedio que mejor represente la posición de la frase. No trate de obtener el mismo número de ítems, o cualquier otra distribución espacial, en el continuo de 1 a 11.

Por favor, antes de empezar, lea una buena cantidad de expresiones para entender el carácter de los ítems.

Queremos insistir en que no se trata de dar una opinión personal de acuerdo o desacuerdo acerca de cada afirmación, sino solamente de estimar su lugar en una escala de 1 a 11 puntos.

[Para el ejemplo presentamos solamente los primeros 29 ítems]

- 1. El modelo desarrollista fortalece la desigualdad social.
- 2. Tratóndose de programas políticos, "mal" vale malo conocido que bueno por conocer".
- 3. Todos los asuntos económicos de interés nacional deben estar a cargo exclusivo del Estado.
- 4. Los trabajadores deben tomar en sus manos la conducción del Estado; el Estado debe tomar en sus manos la conducción de la economía.
- 5. La desigualdad social ha existido siempre, y es necesaria para el desarrollo de la sociedad.
- 6. Los sindicatos deberían limitar sus actividades a las reivindicaciones económicas de sus representados.
- 7. La función del Estado es la de participar dinámicamente en el desarrollo económico, social y político del país.
- 8. La propiedad privada es un derecho natural del hombre y debe ser respetado y mantenido.

9. Las políticas sobre la distribución del ingreso son pura demagogia que sólo propician el enriquecimiento de los políticos.
10. La principal causa de la inflación es el anhelo de los empresarios de aumentar sus ganancias.
11. El Estado debería dejar absolutamente a criterio de los padres de familia, el tipo de educación que prefieren para sus hijos.
12. Todo grupo de interés debe tener igual representación ante el Estado.
13. Los trabajadores producen la riqueza; los patronos se la embolsan.
14. La política del control de la natalidad es una política al servicio del imperialismo.
15. Los que más pierden con la inflación son los empresarios.
16. La importancia creciente del sindicalismo en el país representa un peligro para la democracia.
17. Antes de la Reforma Agraria se producía más y mejor en el campo.
18. La mejor manera de resolver los problemas es encontrar el justo "término medio" y no caer en los extremos.
19. No se trata del sistema tal o cual; para que el país progrese hay que ir bajando más, y punto.
20. El camino de México es el de una economía mixta: el gobierno como promotor, y la iniciativa privada como participante activa en el proceso de desarrollo.
21. El éxito se debe al esfuerzo personal.
22. La nacionalización de la industria minera sólo caería en el burocratismo y la mala administración estatal.
23. La mejor forma de representación ante el Estado es a través de grupos de interés y no a través de partidos políticos.
24. La intervención del gobierno en la economía agrícola sólo ha traído desorden y caos.
25. El deficiente desarrollo agrícola del país se debe a la apatía y flojera de los campesinos.
26. La Revolución Mexicana será verdadera sólo cuando se realice una reforma agraria total.
27. La fuerza política de un grupo debe ser independiente del poder económico de sus miembros.
28. La actividad económica es privativa de los particulares y el Estado debe limitarse a coordinar tal actividad.
29. El marxismo es una doctrina exótica que no toma en cuenta nuestra idiosincrasia ni nuestra tradición.

### 3) Asignación de valores de escala

Vamos a estudiar ahora la distribución de las respuestas (de 1 a 11) en cada ítem, según las respuestas sobre ubicación del ítem dadas por los jueces. Podríamos calcular los valores promedios (media aritmética) y las desviaciones estándar ( $\sigma$ ), para cada ítem. Sin embargo —y aquí de hecho se revela que la escala es más ordinal que intervalar— existe mayor exactitud y representa un método más rápido el calcular valores de mediana ( $Mdn$ ) y de distancia intercuartil ( $Q_2 - Q_1$ ). En las siguientes páginas ilustramos con ejemplos de análisis gráficos. (Ver cuadros 5, 6 y 7.)

En el primer ejemplo (Cuadro 5) hemos obtenido una mediana de 2.3. La mediana indica el "valor" del ítem a lo largo de la variable, es decir que se trata en este caso de un ítem al lado positivo de la escala (muy dogmático). La distancia intercuartil es igual a 2.2. Esta cifra indica la "calidad" del ítem. Si menor la distancia intercuartil, mayor el grado de "calidad" del ítem, es decir, que la adjudicación del valor del ítem por parte de los jueces es similar. Los valores altos de distancia intercuartil, por el contrario, indican diferencias entre los jueces en cuanto a la apreciación sobre el valor adjudicado al ítem.

El ejemplo 2 (Cuadro 6) indica el caso ideal, en el que todos los jueces están de acuerdo en cuanto al valor que debe ser adjudicado al ítem.

El ejemplo 3 (Cuadro 7) representa un ítem de escasa utilidad, ya que como se ve los jueces le adjudicaron valores muy distintos, representando el valor mediano del ítem una adjudicación casi aleatoria de los jueces en cuanto al valor 1, 2, 3... u 11 que se le adjudique al ítem como expresando una posición en el continuo.

En los tres ejemplos presentados se han utilizado 56 jueces, y el método para calcular la mediana y los cuartiles ha sido proyectando líneas sobre la distribución gráfica. A partir de la distribución de frecuencias sobre los valores de ítems (1 a 11), resulta simple calcular la mediana y los cuartiles según las siguientes fórmulas:

$$Mdn = v + i \left( \frac{\frac{N}{2} - F_p}{F_p} \right)$$

Donde:

$v_i$  = Límite exacto inferior del intervalo que contiene la mediana.

$i$  = Amplitud del intervalo de clase (en nuestro caso igual a 1).

$N$  = Número de casos.

$F_p$  = Suma total de las frecuencias inferiores al intervalo que contiene la mediana.

$F_p$  = Frecuencias del intervalo que contiene la mediana.

Para el cálculo de cuartiles:

$$Q_3 = v_i + i \left( \frac{\frac{N}{4} - F_p}{F_p} \right)$$

$$Q_1 = v_i + i \left( \frac{\frac{N}{4} - F_p}{F_p} \right)$$

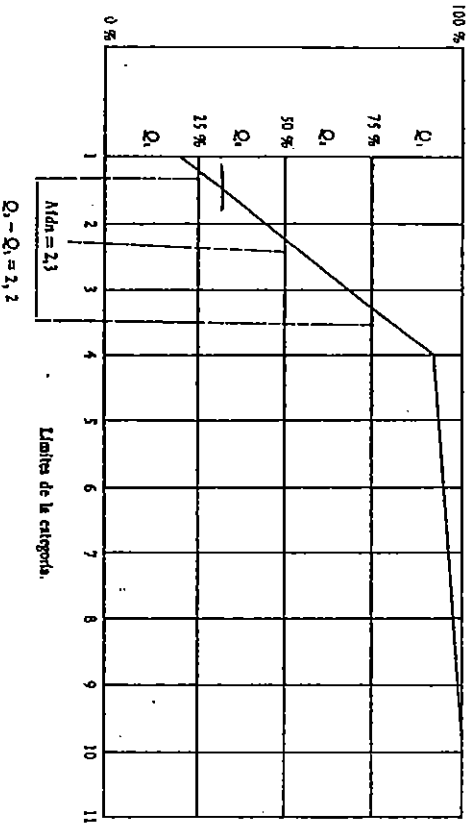
Cuadro 5. Ejemplo 1: Escala de actitud ítem bueno tipo Thurstone mide Dogmatismo

Item núm.  173

Caja núm.	1	2	3	4	5	6	7	8	9	10	11
Frecuencia	////	////	////	////	////	/	//				/
Frecuencia en números	12	11	12	13	4	1	2				1
Frecuencia acumulativa	12	23	35	48	52	53	55				56
Porcentaje acumulativo	21.4	41.1	62.5	85.7	92.5	94.6	98.3				100

56  
ítems

Representación gráfica y cálculo de Mdn y Q



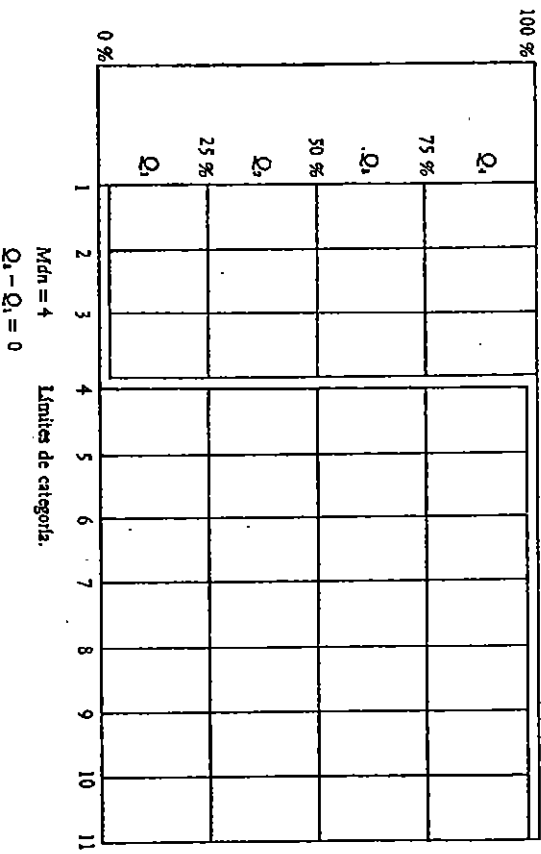
Cuadro 6. Ejemplo 2. Escala de actitud tipo Thurstone. Ítem Ideal.

Item núm.  X

Caja núm.	1	2	3	4	5	6	7	8	9	10	11
Frecuencia											
Frecuencia en números	0	0	0	56	0	0	0	0	0	0	0
Frecuencia acumulativa	0	0	0	56	56	56	56	56	56	56	56
Porcentaje acumulativo	0	0	0	100	100	100	100	100	100	100	100

56  
ítems

Representación gráfica y cálculo de Mdn y A



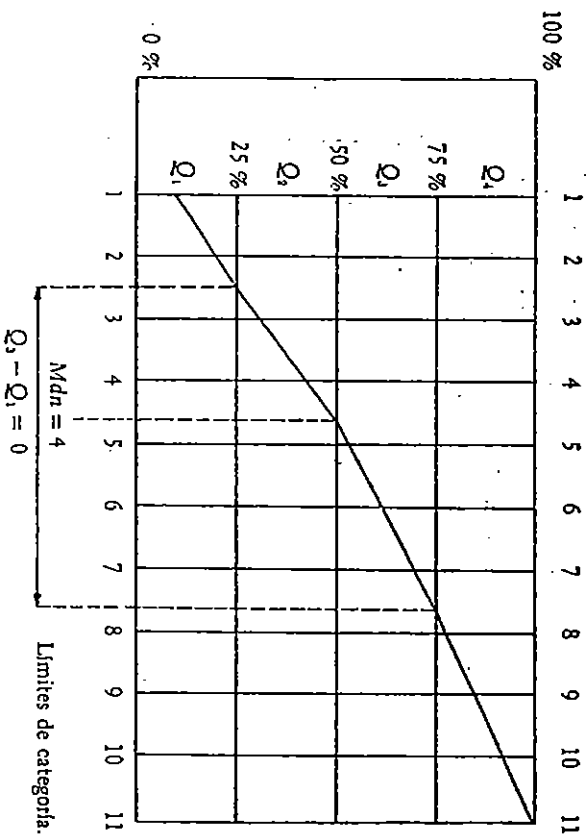
Item n.º	1	2	3	4	5	6	7	8	9	10	11
Caja n.º	1	2	3	4	5	6	7	8	9	10	11
Frecuencia	III	IIII	IIII	IIII	IIII	IIII	IIII	IIII	IIII	IIII	IIII
Frecuencia en números	3	7	8	6	6	4	6	4	5	2	5
Frecuencia acumulada	3	10	18	24	30	34	40	44	49	51	5
Porcentaje acumulativo	5.4	17.8	32.2	42.9	53.6	60.8	71.2	78.5	87.5	91.0	100

Item n.º

172

56 jueces

Representación gráfica y cálculo de Mdn y A.



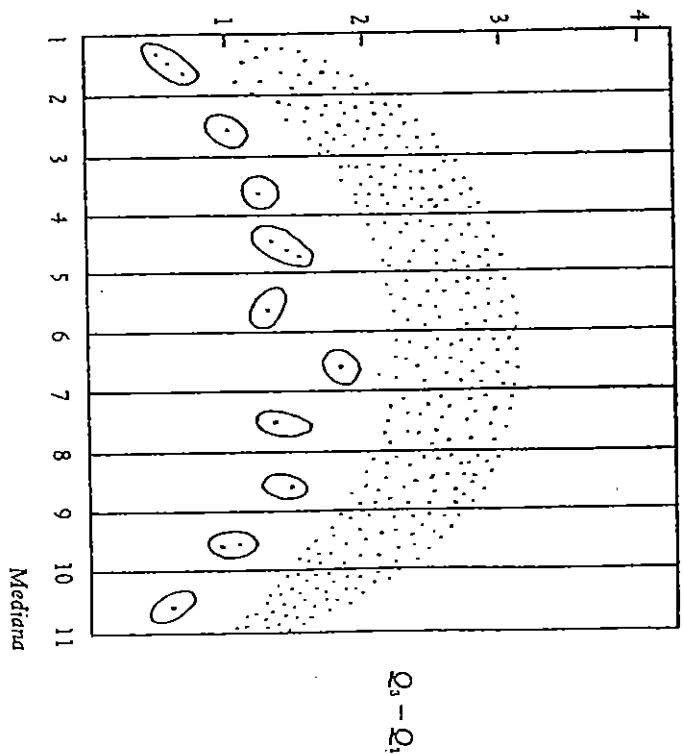
Donde:

- $v_s$  = Límite exacto superior del intervalo que contiene el cuartil.
- $v_i$  = Límite exacto inferior del intervalo que contiene el cuartil.
- $i$  = Amplitud del intervalo de clase.
- $F_a$  = Frecuencias por encima del intervalo que contiene el cuartil.
- $F_d$  = Suma total de frecuencias por debajo del intervalo que contiene el cuartil.
- $F_p$  = Frecuencias en el intervalo que contiene el cuartil.

4) La selección de los ítems

La selección final de los ítems presentados a los jueces se realiza en base a los valores de la mediana, y de amplitud intercuartil. La mediana se utiliza para ubicar el peso del ítem en la escala, eligiendo los ítems que se hallen repartidos uniformemente a lo largo de la misma. Necesitamos aproximadamente 2 ítems para cada intervalo de la escala (1-2; 2-3; 3-4 ... 9-10; 10-11).

Figura 3.





La distancia intercuaril es utilizada para decidir cuáles son los mejores ítems dentro de cada intervalo. La Figura 3 indica una distribución de ítems en la cual en el eje horizontal figuran los valores de mediana; y en el eje vertical los valores de distancia intercuaril. Los ítems seleccionados para la escala final están marcados con círculos y son obviamente aquellos que representando cada uno de los intervalos en la escala, presentan distancias intercuariles mínimas.

Normalmente las distancias intercuariles más grandes se presentan en el centro de la escala. Ocuire muchas veces que la selección de ítems para los valores 5, 6, 7 y 8 se complica en la medida que los jueces tienden a dispersar sus evaluaciones fuertemente en el centro; es decir, esto se va a reflejar en distancias intercuariles bastante grandes. Si es que no se obtienen suficientes ítems en la primera versión de los jueces, el investigador deberá construir más ítems para esta área, que deben ser juzgados nuevamente por los mismos jueces. Conviene por supuesto anticiparse a este tipo de problemas, teniendo especial cuidado cuando se construyen los ítems iniciales.

#### La versión final de la escala

Está constituida por los 15 o 25 ítems seleccionados como confiables. Los ítems son presentados a los sujetos solamente con dos alternativas de respuestas: acuerdo-desacuerdo.

La forma de presentación de la escala es diferente de aquella presentada a los jueces. En un cuestionario, por ejemplo, se eliminan las "cajas", soliciando al sujeto que responda únicamente si está de acuerdo o en desacuerdo con cada uno de los juicios.

A manera de ejemplo, presentamos algunos ítems que han sido contestados por un sujeto hipotético.

1. El éxito se debe al esfuerzo personal.	<input checked="" type="checkbox"/> De acuerdo
	<input type="checkbox"/> En desacuerdo
2. El futuro nos deparará mejores condiciones de vida.	<input type="checkbox"/> De acuerdo
	<input checked="" type="checkbox"/> En desacuerdo
3. El gobierno debería tomar todas las decisiones.	<input type="checkbox"/> De acuerdo
	<input checked="" type="checkbox"/> En desacuerdo
4. Casi nadie considera el trabajo que yo hago.	<input type="checkbox"/> De acuerdo
	<input checked="" type="checkbox"/> En desacuerdo
5. Es bueno que la Iglesia se modere.	<input type="checkbox"/> De acuerdo
	<input checked="" type="checkbox"/> En desacuerdo
6. Yo no podría ser bueno de carácter.	<input checked="" type="checkbox"/> De acuerdo
	<input type="checkbox"/> En desacuerdo
Etcétera.	

#### 5) La adjudicación de puntajes a los sujetos

El investigador tiene que registrar en un código los valores de mediana adjudicados por los jueces a los ítems seleccionados para integrar la versión final de la escala. Los valores intercuariles no se utilizan para el cómputo de los puntajes finales de los sujetos. El puntaje final de un sujeto será entonces simplemente el promedio de los valores de escala de los ítems respondidos en forma afirmativa o "de acuerdo".

Supongamos que nuestro sujeto hipotético haya respondido "de acuerdo" únicamente a los ítems 1 y 6; es decir que en el resto de los ítems su respuesta ha sido "en desacuerdo". Supongamos entonces que los valores de escala dados por los jueces hayan sido:

Ítem	Peso del ítem en la variable
1	1.2
2	4.6
3	7.8
4	3.7
5	10.8
6	1.6
7	9.8
8	5.6
etc.	etc.

Consecuentemente, el puntaje correspondiente a nuestro sujeto será:

$$\frac{1.2 + 1.6}{2} = 1.4$$

Los valores en la escala utilizados como una variable en el análisis de alguna investigación, pueden también ser utilizados de varias formas. Se pue-

Peso en la escala	Núm. de individuos	Frecuencias acumulativas	grupo
Thurstone			bajo
1	34	34	
2	54	88	
3	42	130	
4	39	169	
5	15	184	grupo
6	5	189	alto
7	52	241	
8	84	325	
9	82		
10	72		
11	121		
	600		

den tomar los distintos valores en los sujetos y establecer correlaciones entre la actitud y alguna otra variable, o bien dicotomizar las respuestas de la siguiente manera:

Supongamos que la distribución en nuestra variable entre 600 individuos en una encuesta es como aparece en la página anterior (supongamos que todos han contestado la pregunta).

Lo que hicimos con la distribución fue lo siguiente:

- Contabilizamos el número de sujetos que obtuvieron los diferentes puntajes en la escala; aquí particularmente ubicamos a los sujetos que obtuvieron, por ejemplo, 4,6 en el intervalo 5. Nótese que entonces, la amplitud de cada intervalo de "peso en la escala", para el caso del intervalo 5, va de 4,5 a 5,49.
- Calculamos las frecuencias acumuladas.
- Si buscamos dicotomizar la variable, el estadístico a utilizar sería la mediana, que en nuestro caso tiene un valor igual a 8,2, es decir, cae en la categoría 7,5 a 8,49.
- Procedemos entonces a la dicotomización y obtenemos 2 grupos, a los cuales les llamaremos grupo *alto* y grupo *bajo*.
- Podemos cruzar entonces nuestra variable con alguna otra, tal como figura en el ejemplo,

Variable Y (otra variable cualquiera)

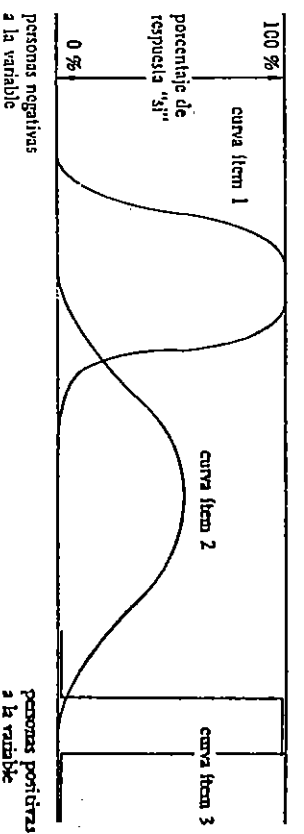
Variable X (Thurstone)	Variable Y			Total
	Alto	Mediano	Bajo	
Grupo alto	15	130	180	325
Grupo bajo	221	52	2	275
	236	182	182	600

### C) Comentarios

#### 1) Items o series de items acumulativos o diferenciados

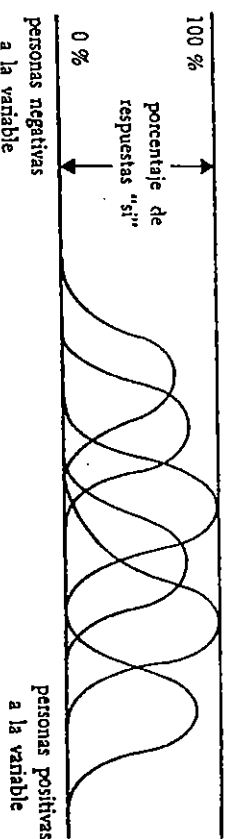
Los *items* de la escala Thurstone son *items* diferenciadores. Presentamos algunos ejemplos de diferentes *items* en una escala Thurstone. (Ver en la página siguiente.)

El eje horizontal representa todas las personas investigadas ordenadas en tal forma que las que detentan una actitud más negativa con respecto a la variable se ubican a la izquierda y las más positivas a la derecha. En el eje vertical está representado el porcentaje de respuestas "de acuerdo" en un *item*.



*Ítem 1*: es un *ítem* que recibe respuestas "No" de las personas que tienen una fuerte posición ya sea positiva o negativa. Un grupo de personas que tengan una determinada posición a lo largo de la escala, responderán este *ítem* con un "Si". En el caso del *ítem 2*, la única diferencia es que solamente el 50% de las personas responderá "Si" al *ítem* en el lugar de la escala donde la probabilidad de alta tasa de respuestas "Si" es más grande. El *ítem 3* es un *ítem* donde el 100% de los respondientes con ciertas posiciones en la escala responderán "Si" y todos los demás "No". Ejemplo: "¿Está usted entre los 178 y los 179 cm de estatura?"

La escala Thurstone en su totalidad forma una serie de *ítems* diferenciadores. La siguiente figura puede dar una idea de la estructura de una escala Thurstone. Hemos reducido el número de *ítems* a 6 para facilitar la lectura de la figura.



#### 2) La variación en el instrumento, los sujetos o ambos

En la escala de Thurstone, solamente el estímulo aislado recibe valor en la escala. El conjunto de tareas para los jueces es escoger los *ítems* dentro de los intervalos de igualdad-aparición (11 cajas) tratando de disminuir cualquier variación debido a su propia posición con respecto a la actitud. Es decir que en el caso de la escala Thurstone tenemos el enfoque que corresponde a la variación centrado en el instrumento. La escala Thurstone entonces pertenece al *test* Tipo B de Coombs.

3) *El problema con las dimensiones*

Una desventaja de la escala de Thurstone es la dificultad de controlar las dimensiones en el variable. Supongamos que una persona ha contestado "de acuerdo" a los siguientes ítems de una escala Thurstone de 30 ítems:

Ítem	Valor
4	5.9
9	6.2
11	6.9
19	7.2
28	7.8
	<hr/>
	34.0

Su promedio es entonces  $\frac{34.0}{5} = 6.8$

Pudo haber recibido esa suma particular mediante una serie de operaciones. Los cinco ítems con los que concordó pudieron haber tenido asignados los números 4.1; 4.2; 8.7; 8.5; 8.9 o por ejemplo: 1.9; 2.2; 7.4; 10.7 y 10.8. En ambos casos el promedio es 6.8. Las posibilidades aumentan todavía más si consideramos que, de hecho, el individuo podía haber concordado con más o menos de 5 ítems. ¿Cuál podría ser la razón de esto? Como simplificación gruesa del problema podemos pensar que la serie 4.1; 4.2; 8.7; 8.5 y 8.9 indica dos dimensiones, una de las cuales aparece alrededor de 4 y la otra alrededor de 9. Aparentemente los dos primeros ítems tienen para el sujeto un significado distinto que los tres siguientes. Para controlar esta desventaja de la escala, Thurstone y Chave (1929) proveyeron a la escala de un criterio burdo basado en las respuestas: el llamado *index of similarity* (índice de similitud).

D) *Ventajas y desventajas de la escala Thurstone*

*Ventajas.* a) La principal ventaja de esta técnica es que permite hacer una distribución de un grupo dado, a lo largo de la actitud que se desea investigar, precisamente porque los ítems fueron diseñados y seleccionados a los efectos de cubrir el continuo. b) Supone una medida más refinada que la escala Lickert, ya que el puntaje de los ítems se deriva de una ponderación basada en el juicio de jueces expertos o al menos informados. c) En la medida que la escala final contiene más ítems que la de Lickert, es más confiable que ésta. d) Si es tratada como escala intervalar, permite comparar puntuaciones y cambios de actitud en los sujetos.

*Desventajas.* a) Su elaboración es larga y compleja. b) Pese a su tratamiento como escala intervalar, su verdadera naturaleza en términos de nivel de medición corresponde al nivel ordinal. c) Fácilmente se introducen otras dimensiones distintas de las que se quiere medir. d) Discrimina poco en los extremos de la distribución. e) Los ítems neutrales carecen de significado

mayor, y a menudo se ubican en esta posición ítems que no se refieren a la dimensión tratada. f) Distintas configuraciones de respuesta resultan en el mismo puntaje final. g) Los jueces pueden introducir sesgos difíciles de detectar.

## LA ESCALA GUTTMAN

Una de las desventajas mayores en las dos escalas que examinamos hasta ahora—Lickert y Thurstone—era que ninguna de ellas garantizaba que el instrumento mida una dimensión única.

La escala Guttman, conocida como método del escalograma o análisis de escalograma, soluciona el problema de la unidimensionalidad. Su objetivo es el de definir lo más claramente posible qué es lo que está midiendo la escala, entendido esto como un problema de unidimensionalidad. Por el tipo especial de tratamiento al que se somete a la escala se busca la eliminación de factores extraños a la característica o dimensión que se pretende medir.

La escala Guttman es de tipo acumulativo, ya que la respuesta positiva a un ítem supone que los ítems anteriores también han sido respondidos en forma positiva. Se busca pues una coherencia en las pautas de respuesta de los sujetos, y esa coherencia es garantizada por medio de un *coeficiente de reproducibilidad*. El tamaño del coeficiente (valor máximo 1.00) señala el grado por el cual la escala es acumulativamente perfecta o casi perfecta. En una escala cuya reproducibilidad es perfecta, las respuestas de los sujetos a todos los ítems pueden ser reproducidas por el solo conocimiento de su posición de rango.

Veremos más adelante que además de la reproducibilidad, hay que tomar en cuenta otros factores, tales como el alcance de la distribución marginal, la pauta de los errores, el número de ítems en la escala y el número de categorías de respuesta.

El método de escalograma de Guttman combina aspectos de construcción utilizados en las escalas Lickert y Thurstone, a más de los distintos cálculos de los coeficientes mencionados en el párrafo anterior, en razón de que utiliza 2 técnicas: a) Siguiendo los procedimientos de la escala Lickert; y b) La técnica de la escala discriminatoria de Edwards y Kilpatrick. Vamos a desarrollar primero los procedimientos siguiendo las técnicas de Lickert y luego utilizaremos la técnica de Edwards y Kilpatrick.

A) *La construcción de un escalograma Guttman*

Los pasos a dar en la construcción de una escala de este tipo son: 1) Se construye una serie de ítems relevantes a la actitud que se quiere medir. 2) Se administran los ítems a una muestra de sujetos que van a actuar como jueces. 3) Se asignan puntajes a los ítems según la dirección positiva o negativa del ítem. 4) Análisis de ítems para la formación de series acumulativas. 5) En base a los ítems seleccionados se construye la escala final.

Para los pasos 1, 2 y 3 se siguen los procedimientos señalados en la escala

Lickert para la construcción de *ítems*, la administración a los jueces y la asignación de puntajes (ver escala Lickert).

4) El análisis de los ítems

Una vez aplicados los *ítems* a los jueces, procedemos al análisis de los *ítems* en su conjunto. La idea es formar una serie acumulativa de *ítems*. Para ello, vamos a diseñar un escalograma.

a) Primero computamos el puntaje total para cada uno de los jueces (es decir sumamos los valores obtenidos en cada uno de los *ítems*).

b) Ordenamos a los jueces según el puntaje total, desde el puntaje más alto al puntaje más bajo.

Si nuestra serie de *ítems* fuese perfecta, todas las celdas cruzadas (*crossed cells*) en el escalograma estarían en una posición sobre una diagonal que corre desde el ángulo superior izquierdo hasta el ángulo inferior derecho. Mientras más alto sea el número de desviaciones a esta diagonal, más baja será la reproducibilidad y menos idéntica a una serie de *ítems* acumulativos será nuestra serie.

Antes de entrar a un análisis en detalle del Cuadro 8, vamos a mostrar un escalograma más simple con el fin de clarificar la idea de *acumulación* y *de error*.

Supongamos que 20 jueces hayan respondido a 6 *ítems* en términos de "acuerdo" o "desacuerdo". Ordenamos a los jueces, figurando en el Cuadro 8 los "acuerdos" con una x, y los "desacuerdos" con una O. La distribución se muestra en el Cuadro 8.

El Cuadro 8 muestra una serie de *ítems* escalonados. El *ítem* 4 ocupa la primera posición en la escala en razón de que se dieron en él 3 respuestas positivas (de acuerdo); el *ítem* 2 ocupa el segundo lugar ya que se dieron 6 respuestas positivas, y así sucesivamente hasta llegar al último *ítem*, el número 3, con el que 16 de los 20 jueces acordaron.

En términos de escalograma, las respuestas redondeadas con un círculo son "errores", esto es, respuestas que caen fuera de la pauta general del escalograma señalado con la línea segmentada.

El cuadro analítico para el Cuadro 8, sería:

$$\begin{aligned} \text{Número de preguntas:} & \quad 6 \\ \text{Número de jueces:} & \quad 20 \\ \text{Cantidad total de respuestas:} & \quad 6 \times 20 = 120 \\ \text{Cantidad de errores:} & \quad 7 \\ \text{Reproducibilidad} & = 1 - \left( \frac{\text{Cantidad de errores}}{\text{Cantidad total de respuestas}} \right) = 0.94 \end{aligned}$$

Compliquemos un poco el ejemplo, utilizando 5 categorías o alternativas de cada *ítem*.

Supongamos que 30 jueces han juzgado 6 *ítems* en una escala de 5 puntos:

- 4— Muy de acuerdo
- 3— De acuerdo
- 2— Indeciso
- 1— En desacuerdo
- 0— Muy en desacuerdo

Cuadro 8. Análisis de escalograma. Respuestas de 20 jueces a 6 ítems en términos de acuerdo-desacuerdo

Rango del juez	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Puntaje
1	X																				6
2	X	X																			6
3	X	X	X																		6
4			X	X																	5
5				X	X																4
6					X	X															4
7						X	X														3
8							X	X													3
9								X	X												3
10									X	X											3
11										X	X										3
12											X	X									3
13												X	X								3
14													X	X							2
15														X	X						1
16															X	X					1
17																X	X				1
18																	X	X			1
19																		X	X		0
20																			X	X	0



La distribución de los puntajes en los ítems adquiere ahora la siguiente forma:

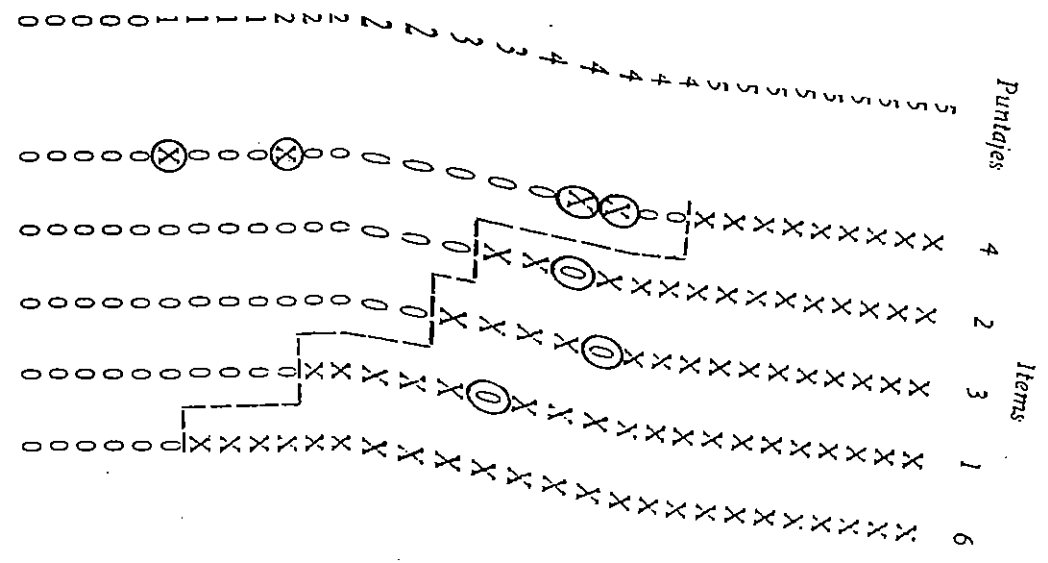
Item 4 tiene algunos errores mientras que el 5, a pesar de la dicotomización contiene más errores que los demás. La distribución de los puntajes en el ítem 1 indica la necesidad de unir las nuevas categorías 2 y 1 con el fin de

CUADRO 10

Puntajes	Item 1		Item 2		Item 3		Item 4		Item 5		Item 6	
	2	1	2	0	2	0	2	0	2	0	2	0
12	X											
12	X											
12	X											
11	X											
11	X											
11	X											
10	X											
10	X											
10	X											
9	X											
9	X											
8	X											
7	X											
6	X											
5	X											
5	X											
5	X											
4	X											
4	X											
4	X											
4	X											
2	X											
2	X											
2	X											
0	X											
0	X											
0	X											
0	X											

Obsérvese que con la nueva clasificación de los ítems, el ítem 6 no contiene ahora ningún error, siendo a su vez el que menos discrimina entre los puntajes extremos. Los ítems 2 y 3 escalan bastante bien y con poco error. El

CUADRO 11



bilidad será igual a:

$$r_p = 1 - \frac{\text{Total de errores}}{\text{Total de respuestas}} = 1 - \frac{38}{990} = .913$$

Valor satisfactorio, sin embargo, si unimos las categorías 1 y 2 del ítem 1, y eliminamos el ítem 5; detengamos entonces el escalograma de la página anterior, en el cual los ítems aparecen ordenados por grado de dificultad. Puede verse ahora que los errores han sido reducidos en forma considerable, aunque a costa de reducir la discriminación al interior de los ítems, más de eliminar el ítem originalmente designado con el número 5. El lector notará que con los nuevos pesos asignados a los ítems se produce un reordenamiento en el rango de los ítems como consecuencia de la alteración en los puntajes totales. Las líneas segmentadas representan los puntos de cortes.

### B) La determinación de los errores

Hay dos métodos que tienden a dar resultados diferentes: la técnica de Cornell y la técnica de Goodenough.

Comencemos con la técnica de Cornell, la cual se realiza estableciendo puntos de separación en el orden de rango de los jueces, puntos que los separan en distintas categorías, tales como se definirían de ser la escala perfecta. Ésta es la técnica utilizada cuando señalamos puntos de separación por medio de líneas segmentadas. Si observamos el Cuadro 10, vemos que el ítem 1 necesita dos puntos de separación: el primero cae entre la última persona con puntaje 10 y la primera persona con puntaje 9. Todos los sujetos o jueces por encima de este punto de separación deberían tener puntaje 2 en el ítem. En términos de errores encontramos pues 5 errores (cuatro por encima y uno por debajo del punto de separación). El siguiente punto de separación lo establecemos entre la última persona con puntaje 5 y la primera persona con puntaje 4. No existe ningún error.

Finalmente, en el último punto de corte que corresponde al puntaje 0 tenemos 1 error. Así pues, el ítem 1 tiene un total de 6 errores. Dado que hay 80 jueces y 3 categorías, el porcentaje de error es de 6.7%. Cuando el error es mayor que el 10% y el ítem permite reagrupación es necesario reevaluarla, o de lo contrario conviene eliminar al ítem. Se opera de la misma manera con los otros ítems.

Para el cálculo del coeficiente de reproducibilidad por la Técnica de Cornell, hay que seguir entonces los siguientes pasos: a) Determinar los cutting points de manera tal de minimizar los errores. b) Ningún cutting point debe estar tan por encima o tan por debajo que la categoría menor incluya menos del 10% de los jueces. c) Ningún cutting point debe ser realizado de manera tal que el número de errores sea mayor que el número de jueces en la categoría.

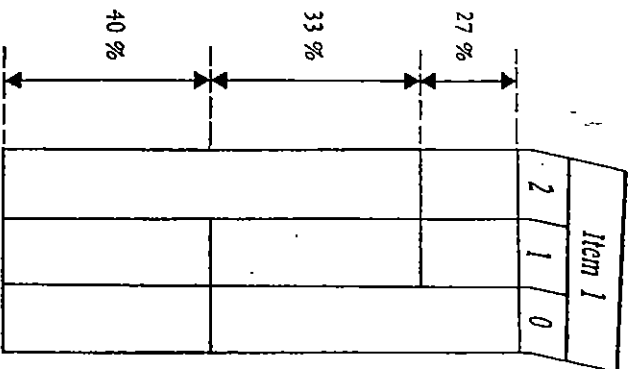
### PUALAS PARA LA MEDICION DE ACTITUDES

Para controlar si el lector entendió la manera de calcular el coeficiente de reproducibilidad por la técnica de Cornell, lo invitamos a hacer el cálculo para el Cuadro 10 por su cuenta. Para los cutting points indicados en el cuadro, el coeficiente debe ser igual a .949. De hecho, habría que aclarar que el ítem 5 debería ser eliminado, con lo cual el coeficiente de reproducibilidad aumentaría a un valor aceptable, tal como vimos en el Cuadro 11. La técnica de Goodenough, es un poco más complicada y se basa en el cálculo de errores en base a pautas marginales.

Los cutting points en esta técnica se determinan según la distribución de las respuestas de los jueces a cada una de las alternativas en los distintos ítems. Los cálculos para el Cuadro 10 nos darían los siguientes valores:

	Ítems					
Categorías	1	2	3	4	5	6
Frecuencias	2	1	0	2	0	2
Porcentajes	8	10	12	14	16	15
	15	15	15	14	16	16
	16	16	14	16	16	14
	24	24	6			
	27	33	40	47	53	50
	50	50	50	47	53	47
	53	47	53	47	53	80
	20					

Los cutting points para cada uno de los ítems deben seguir entonces los porcentajes respectivos. En el caso del ítem 1, los cortes serían entonces:



Es decir, que en el ítem 1, el primer corte o *cutting point* caería entre el primer sujeto con puntaje total 10 y el segundo con puntaje 10, y así sucesivamente para el resto de los ítems.

Mediante estos *cutting points*, se van a determinar pautas de respuestas correspondientes a cada corte. La pauta de respuesta es la manera "correcta" en que deberían distribuirse para cada puntaje total de cada juez si la escala tuviera perfecta escalabilidad. Cada respuesta que no sigue la pauta de respuesta ideal se considera un "error". En nuestro ejemplo, un juez con puntaje de 8 debe seguir una pauta: 0-0-2-2-2-2. Un sujeto con puntaje 6, debe consecuentemente tener una pauta de respuesta 0-0-0-2-2-2. Es decir que primero hay que ordenar a los ítems en términos de escalabilidad, para luego determinar las pautas correspondientes a cada valor. El principio de la técnica es bastante simple. Para 4 ítems dicotomizados y con valores de 1 y 0 para alternativas de respuesta "de acuerdo" y "en desacuerdo", respectivamente, un puntaje total de 3 puede seguir cuatro tipos de pautas diferentes:

- a) 0 1 1 1
- b) 1 0 1 1
- c) 1 1 0 1
- d) 1 1 1 0

De estas pautas, solamente la primera es correcta, para ítems que están ordenados en forma acumulativa. Cada una de las pautas b), c) y d) tiene respectivamente 2 errores (uno por tener un 1 donde debería haber un 0, y otro por tener un cero donde debería tener un 1).

		Ítems				Puntaje total	Pauta de respuesta			Errores
1	2	3	4	1	1		1	1		
1	0	1	0	1	0	4	1	1	1	0
X	X	X	X	X	X	3	1	1	1	0
X	X	X	X	X	X	3	0	1	1	2
X	X	X	X	X	X	3	0	1	1	2
X	X	X	X	X	X	3	0	1	1	2
X	X	X	X	X	X	2	0	1	1	2
X	X	X	X	X	X	2	0	1	1	2
X	X	X	X	X	X	2	0	1	1	2
X	X	X	X	X	X	1	0	0	1	0
X	X	X	X	X	X	0	0	0	0	0
X	X	X	X	X	X	0	0	0	0	0
Total errores										6

Para hacer más claro el método, en la página anterior dimos un ejemplo más simple, en el que presentamos 4 ítems dicotomizados, las pautas de respuesta y el cálculo de los errores.

El coeficiente de reproducibilidad, se calcula según la fórmula ya conocida de:

$$r_p = 1 - \frac{\text{Cantidad de errores}}{\text{Número total de respuestas}}$$

En el ejemplo arriba citado:

$$r_p = 1 - \frac{6}{60} = .90$$

El coeficiente de reproducibilidad nos indica la proporción de respuestas a los ítems que pueden ser correctamente reproducidas.

#### B) Ejemplos

Veamos ahora dos ejemplos, a partir de los cuales presentaremos criterios adicionales al coeficiente de reproducibilidad para la determinación del universo de contenido en la escala Guttman.

Cuadro de la página siguiente: escalograma Guttman para la medición de nivel de vida determinada en función de materiales empleados en la construcción de viviendas y de las instalaciones sanitarias. Muestra de población de Colombres, en el departamento de Santa Cruz en Tucumán, Argentina. Respuestas de 38 jefes de familia (Arg. Hernández, c/s; Universidad de Tucumán, Argentina, 1956).

#### REFERENCIAS

Ejemplo B	Ejemplo A
Ítem 1: Alejamiento de aguas servidas a pozo ciego y fosa séptica.	Ítem 1: Pileta de lavado.
Ítem 2: Agua corriente.	Ítem 2: Inodoro.
Ítem 3: Eliminador de residuos.	Ítem 3: Ducha.
Ítem 4: Techo de cinc, loza o teja con aislación.	Ítem 4: Revestimiento de cemento o superior en baño o letrina.
Ítem 5: Piso de cemento o mejor.	
Ítem 6: Luz eléctrica.	



Sujetos	Ejemplo A: Vivienda					Ejemplo B: Instalaciones sanitarias					
	Items					Items					
	1	2	3	4	5	6	Sujetos	1	2	3	4
1	X	X	X	X	X	X	3	X	X	X	X
33	X	X	X	X	X	X	15	X	X	X	X
31	X	X	X	X	X	X	17	X	X	X	X
22	X	X	X	X	X	X	7	X	X	X	X
19	X	X	X	X	X	X	11	X	X	X	X
16	X	X	X	X	X	X	30	X	X	X	X
4	X	X	X	X	X	X	2	X	X	X	X
6	X	X	X	X	X	X	1	X	X	X	X
3	X	X	X	X	X	X	4	X	X	X	X
15	X	X	X	X	X	X	5	X	X	X	X
17	X	X	X	X	X	X	8	X	X	X	X
14	X	X	X	X	X	X	6	X	X	X	X
7	X	X	X	X	X	X	33	X	X	X	X
2	X	X	X	X	X	X	31	X	X	X	X
18	X	X	X	X	X	X	22	X	X	X	X
5	X	X	X	X	X	X	19	X	X	X	X
8	X	X	X	X	X	X	16	X	X	X	X
11	X	X	X	X	X	X	14	X	X	X	X
30	X	X	X	X	X	X	9	X	X	X	X
29	X	X	X	X	X	X	18	X	X	X	X
9	X	X	X	X	X	X	13	X	X	X	X
13	X	X	X	X	X	X	25	X	X	X	X
10	X	X	X	X	X	X	24	X	X	X	X
26	X	X	X	X	X	X	29	X	X	X	X
28	X	X	X	X	X	X	27	X	X	X	X
24	X	X	X	X	X	X	26	X	X	X	X
27	X	X	X	X	X	X	32	X	X	X	X
20	X	X	X	X	X	X	12	X	X	X	X
21	X	X	X	X	X	X	20	X	X	X	X
25	X	X	X	X	X	X	28	X	X	X	X
12	X	X	X	X	X	X	23	X	X	X	X
23	X	X	X	X	X	X	10	X	X	X	X
32	X	X	X	X	X	X	21	X	X	X	X

Cálculo Universo de Contenido para el Cuadro 11:

Coefficiente de reproducibilidad:	Ejemplo A	Ejemplo B
Rango marginal mínimo:	.949	1.00
Alcance de distribución marginal:	.72	.60
	.229	.40

El coeficiente de reproducibilidad en el Ejemplo B, cuyo valor es 1.00, nos permite decir, sin tener en cuenta la distribución gráfica, que el sujeto 18 tiene en su casa únicamente revestimiento de cemento (o mejor, en su baño), mientras que el sujeto 33, cuyo puntaje es 3, tiene revestimiento, inodoro y ducha, pero no piletta de lavado.

Es muy difícil lograr escalabilidad perfecta, y consecuentemente existen errores que van a ser interpretados como errores de reproducibilidad. Guttman aconseja que los coeficientes de reproducibilidad no sean menores de .90.

El coeficiente de reproducibilidad ( $r_p$ ) es un criterio necesario, pero no suficiente para la determinación de la escalabilidad de los ítems. Deben tomarse en cuenta otros factores. Stouffer *et al.*<sup>9</sup> señalan cuatro criterios adicionales:

1) Alcance de la distribución marginal; 2) Pauta de errores; 3) Número de ítems en la escala, y 4) Número de categorías de respuestas.

#### 1) Alcance de la distribución marginal

Es el más importante de los criterios adicionales, y debe acompañar al coeficiente de reproducibilidad. El criterio de distribución marginal es determinado por el Rango Marginal Mínimo (m. m. r.) que consiste en el  $r_p$  menos el promedio de los modos de las frecuencias relativas de las distribuciones de los ítems: ( $r_p - MMR$ ).

Para algunos, los valores de este criterio adicional deben variar entre .15 y .35; para otros el mínimo debe ser mayor que .10. Estos valores indican la escalabilidad de los ítems, dato que no es proporcionado por el  $r_p$  de manera completa (es decir, es posible alcanzar valores altos de  $r_p$ —digamos .90—y resultar una escalabilidad inaceptable. Éste es el caso en el cual los *cutting points* están muy próximos entre sí, con el resultado de discriminar solamente en los extremos de la escala y no a lo largo de la misma. En nuestro ejemplo, los valores de  $r_p$  son altos y muy aceptables; los alcances de la distribución marginal, en cambio, son aceptables para el ejemplo A, y demasiado altos para el ejemplo B.

#### 2) Pauta de errores

Cuando el  $r_p$  es menor que .90, pero es escalable, es decir que tiene un  $r_p$  m. m. r. mayor que .10 estamos en presencia de más de una variable; mejor dicho, de una variable dominante y de otra u otras menores, en el área a través de la cual se ordenan los sujetos. Este tipo de escalograma es denominado cuasi-escala. Éste no es el caso de los dos ejemplos que presentamos.

#### 3) Número de ítems en la escala

A mayor número de ítems, mayor la seguridad de que el universo (del cual están ítems son una muestra), es escalable. Es por esto que cuando los ítems están dicotomizados, como es el caso en nuestros ejemplos, es aconsejable que

<sup>9</sup> Stouffer, S. *et al.*: *Measurement and Prediction*. Studies in Social Psychology in World War II, vol. IV, J. Wiley & Sons, N. Y., 1966.

su número sea mayor que 10. Pero puede usarse un número menor de ítems si las frecuencias marginales se colocan en un rango con recorridos del 30% al 70%.

En los ejemplos dados por nosotros el rango de frecuencias es:

Ejemplo A		Ejemplo B	
Item 1 .....	24%	Item 1 .....	30%
Item 2 .....	57%	Item 2 .....	55%
Item 3 .....	60%	Item 3 .....	55%
Item 4 .....	69%	Item 4 .....	60%
Item 5 .....	78%		
Item 6 .....	87%		

De acuerdo al requisito citado más arriba tenemos así alguna seguridad, de que el universo se comporta como la muestra.

4) Número de categorías de respuestas

Es otro criterio para asegurar la escalabilidad; cuanto mayor el número de categorías, mayor la seguridad de que el universo es escalable. Por ello, a pesar de la necesidad de reducir las categorías por razones prácticas (disminución del número de errores), hay que asegurarse de que tal reducción no es la resultante de obtener frecuencias marginales extremas (.90-.10) que, como vimos más arriba, no permiten errores, pero artificialmente.

Si mantenemos el número de alternativas de respuestas, a pesar de que aumentará el número de errores, disminuimos la posibilidad de que aparezca una pauta escalable cuando de hecho el universo no lo es.

C) La técnica de la escala discriminativa de Edwards y Kilpatrick

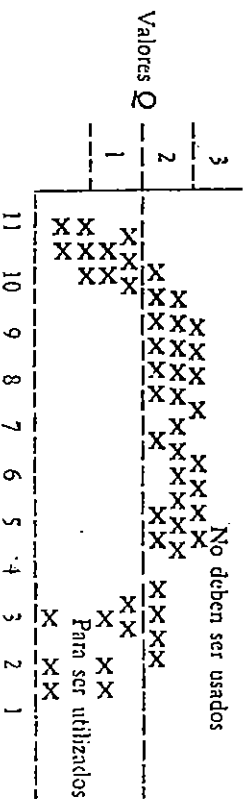
La selección de los ítems iniciales para una escala Guttman ha sido siempre un misterio. En su primer libro Guttman afirma que la selección es cuestión de intuición. Pero, ¿cómo opera esta intuición? ¿No existen reglas o hechos que puedan ser usados cuando se seleccionan los ítems escalables? Apparently la intuición de Guttman opera satisfactoriamente porque muchas escalas que han usado esta "técnica" terminan por ser perfectamente escalables. Y aun en las publicaciones más tardías la vía de la intuición parece ser la más correcta.

Si nosotros recapitulamos el proceso total para el investigador, vemos que él, inicialmente, tiene un número de afirmaciones que más o menos miden adecuadamente la variable. De éstos seleccionará un pequeño número al cual aplicará el análisis de escalograma o tal vez la técnica H. Nada indica en la técnica de Guttman que los ítems deben ser elegidos de manera que representen diferentes pasos en el continuo psicológico. Como una cuestión de hecho, todos podrían tener una posición, digamos de 7 en un continuo tipo Thurstone. El único criterio es que los ítems sean escalables de acuerdo con los efectos discriminativos entre un grupo alto y un grupo bajo.

Para superar esto, se distribuyen representativamente los ítems a lo largo del continuo psicológico, y para establecer algunas reglas para la selección de los ítems iniciales de una escala Guttman, Edwards y Kilpatrick desarrollaron una técnica llamada: técnica de la escala discriminativa (*the scale discrimination technique*).

Los principales pasos en el procedimiento podrían ser expuestos de la siguiente manera:

- 1) Por medio de la técnica de Thurstone de los intervalos aparentemente iguales (*the Thurstone equal appearing interval technique*) un gran número de ítems son colocados a lo largo del continuo.
- 2) Se calcula para cada ítem la posición mediana y los valores Q.
- 3) Para disminuir el número de ítems se seleccionan aquellos que en los diferentes intervalos de la escala tienen valores bajos de Q; para esto deben colocarse a todos los ítems en una figura como la siguiente:



Se coloca el *cutting point* como lo indica la línea — (mediana de la escala Q) y se incluyen aquellos ítems que tienen valores bajos de Q; los que tienen valores altos se excluyen.

- 4) Ahora se vuelcan los ítems que quedan a una escala Likert, adicionando 5 o 6 alternativas de respuesta para cada uno.
- 5) Se aplica la nueva escala a por lo menos 100 jueces, y se separan el grupo alto (Q) y el grupo bajo (Q<sub>b</sub>), de acuerdo con los puntajes totales.
- 6) Para probar el poder discriminativo de los ítems se dicotomizan los pesos como en el siguiente ejemplo:

	Grupo bajo	Grupo alto
(5)	2	4
(4)	8 a	3 d
(3)	7	1
(2)	28	32
(1)	32 c	38 b
(0)	19	17



## D) La técnica H para mejorar la escalabilidad de la escala Guttman

Para mejorar la escalabilidad de la escala Guttman se puede utilizar la técnica de H. La idea es formar por intermedio de esta técnica nuevos *ítems* (*contrived items*) agrupados con algunos de los *ítems* iniciales.

Supongamos que tenemos los datos que aparecen en el cuadro 12 (p. 211). Es decir, 25 individuos han respondido a 10 *ítems*. Nosotros queremos aumentar la escalabilidad formando "*contrived items*".

1) Existen cinco categorías de respuestas para cada *ítem* arbitrariamente numeradas 0, 1, 2, 3, 4.

2) Para cada persona se computa un puntaje total basado en todos los *ítems* que serán sometidos al análisis de escala.

3) Se obtienen las tablas de correlación de cada uno de los *ítems* considerados con el puntaje total provisorio. Para cada *ítem* se hacen distintas participaciones, considerando como categorías positivas las respuestas números 1, 2, 3 y 4, después la 2, 3 y 4, después la 3 y 4 y luego la 4. Empezamos por ejemplo con *ítem* 1, tomando 3 y 4 como positivo:

Cuadro 13

Puntaje total clasificado	0,1,2	Positivos 3,4
35-40	1	6
30-34	4	6
25-29	3	1
20-24	1	
15-19		1
10-14	2	
5-9		
0-4		

4) Seleccionamos los *cutting points* para cada *ítem* que se correlacionan con el puntaje (*score*) total, los cuales son lo suficientemente altos para formar una tabla de 2 por 2 en la cual ninguna celda error tiene una frecuencia mayor que la menor frecuencia de las dos celdas de la diagonal principal. Es decir, en este caso:

Cuadro 14

	0, 1, 2	llegamos a: 3, 4	
35-40 30-34	a) 5	b) 12	Frecuencia Positiva = 12
25-29 20-24	c) 6	d) 2	
15-19 10-14			
5-9 0-4			

Calculamos el  $r_q$ :

$$r_q = \frac{bc - ad}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \frac{12 \cdot 6 - 5 \cdot 2}{\sqrt{17 \cdot 14 \cdot 11 \cdot 8}} = .43$$

Se siguen haciendo los mismos cálculos de  $r_q$  para el *ítem* 1, pero ahora agrupando las categorías de otra manera. Las otras posibilidades son:

	Positivos
0	1, 2, 3, 4
0, 1	2, 3, 4
0, 1, 2	3, 4
0, 1, 2, 3	4

(menos utilizado)

5) Ordenamos todos los *cutting points* de la frecuencia positiva más alta a la más baja. Véase Cuadro 15.

6) Seleccionamos conjuntos triples de *ítems* para constituir los cuatro nuevos *contrived items*. El objetivo principal es seleccionar *ítems* aceptables con la misma frecuencia aproximada para cada triple y espaciar cuanto sea posible estos conjuntos de triples tan extendidos (Cuadro 15).

Cuadro 15.

Item	Categorías Positivas	Frecuencia	Q
6	1234	25	.04
6	234	25	.04
7	1234	24	.08
7	1234	24	.08
8	1234	24	.08
8	234	24	.08
10	1234	24	.08
10	234	24	.08
2	1234	24	.08
2	1234	24	.08
3	1234	24	.08
3	1234	24	.08
1	1234	23	.44
1	1234	23	.44
4	1234	23	.44
4	234	23	.44
7	34	23	.44
8	34	23	.44
9	1234	23	.04
9	234	23	.04
9	34	23	.04
10	34	22	.62
10	34	22	.62
2	234	21	.73
2	34	21	.20
6	34	21	.20
6	34	21	.20
7	34	21	.73
7	34	21	.73
9	34	21	.60
9	34	21	.60
2	34	20	.80
2	34	20	.80
3	234	20	.85
3	234	20	.85
5	1234	20	.80
5	1234	20	.80
9	4	19	.79
9	34	17	.92
3	34	17	.92
5	234	17	.88
5	234	17	.88
1	234	16	.89
1	234	16	.89
6	4	16	.77
6	4	16	.77
8	4	16	.53
8	4	16	.53
4	234	15	.81
4	234	15	.81
1	34	12	.43
1	34	12	.43
7	4	13	.63
7	4	13	.63
2	4	12	.97
2	4	12	.97
4	34	12	.83
4	34	12	.83
5	34	12	.94
5	34	12	.94
3	4	10	.26
3	4	10	.26
5	4	5	.81
5	4	5	.81
1	4	3	.53
1	4	3	.53
4	4	2	.37
4	4	2	.37

} Contrived Item I  
 } Contrived Item II  
 } Contrived Item III  
 } Contrived Item IV

Reglas: a) Si es posible, no repetir los ítems iniciales en diferentes *contrived* ítems. b) Buscar conjuntos de 3 ítems con un máximo de correlación ( $r_q$ ).  
 7) Adjudicar puntajes a cada individuo en cada *contrived item* asignándole un puntaje positivo si él fue positivo en dos o tres de los ítems componentes del *contrived item*.

Cuadro 16. *Contrived items*

Individuos	I	II	III	IV
1	-	-	-	+
2	-	+	+	+
3	-	+	+	+
4	-	-	+	+
5	-	-	+	+
6	-	-	+	+
7	-	+	+	+
8	-	-	+	+
9	-	-	+	+
10	+	+	+	+
11	+	-	+	+
12	-	+	+	+
13	-	+	+	+
14	-	+	+	+
15	+	+	+	+
16	-	-	-	-
17	-	-	-	-
18	-	+	+	+
19	-	+	+	+
20	-	+	+	+
21	-	+	+	+
22	-	+	+	+
23	-	+	+	+
24	-	+	+	+
25	-	+	+	+

• No escalable  
 CR = .96  
 MMR = .76

Entonces hemos llegado a 4 nuevos ítems (*contrived items*) que muestran un mayor coeficiente de reproducibilidad y una mayor diferencia entre-OR y MMR.

## E) La versión final de la escala

Hemos analizado el conjunto total de ítems, llegando finalmente a 10 ítems escalables según los criterios de la escala Guttman. Reproducimos aquí 4 de ellos y la manera de presentarlos en un cuestionario (en el ejemplo que citamos los 10 ítems han sido escalables para 5 valores en las alternativas de respuestas).

1. El patriotismo y la lealtad son los requisitos primarios y más importantes que debe llevar todo buen ciudadano. Ud. está:	<p>Totamente de acuerdo.</p> <p>De acuerdo en general.</p> <p>Ni de acuerdo ni en desacuerdo.</p> <p>En desacuerdo en general.</p> <p>Totamente en desacuerdo.</p>
2. Quien no quiere pelear por su país merece algo mucho peor que la cárcel o los trabajos forzados. Ud. está:	<p>Totamente de acuerdo.</p> <p>De acuerdo en general.</p> <p>Ni de acuerdo ni en desacuerdo.</p> <p>En desacuerdo en general.</p> <p>Totamente en desacuerdo.</p>
3. Cualquier esfuerzo hecho en el entrenamiento militar de Chile, es compensado por los beneficios de su seguridad. Ud. está:	<p>Totamente de acuerdo.</p> <p>De acuerdo en general.</p> <p>Ni de acuerdo ni en desacuerdo.</p> <p>En desacuerdo en general.</p> <p>Totamente en desacuerdo.</p>
4. Para Chile sería un grave error permitir el ingreso de los extranjeros que quitan las oportunidades de trabajo a los nacionales. Ud. está:	<p>Totamente de acuerdo.</p> <p>De acuerdo en general.</p> <p>Ni de acuerdo ni en desacuerdo.</p> <p>En desacuerdo en general.</p> <p>Totamente en desacuerdo.</p>

La escala es unidimensional, es decir, que los resultados obtenidos de los sujetos van a permitir ubicarlos en el continuo favorable-desfavorable a la actitud, en un rango que va (en nuestro ejemplo) de 0 a 40 puntos.

## F) Ventajas y desventajas de la escala Guttman

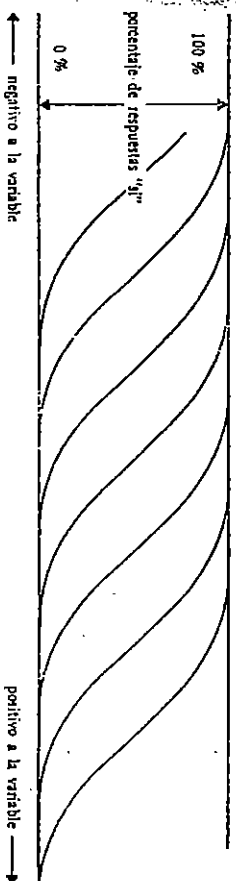
**Ventajas.** a) Se asegura en forma casi definitiva la unidimensionalidad de la escala. b) Conociendo el puntaje de un individuo se puede saber qué grado de acuerdo tuvo con los ítems y ubicarlo así en el continuo de la escala.

**Desventajas:** a) Cuando se trata de medir actitudes complejas conviene hacer un tipo de escala para cada dimensión de la actitud. b) Involucra mayor cantidad de trabajo que en las dos escalas mencionadas anteriormente (Lickert y Thurstone). c) Escalas de este tipo pueden resultar unidimensionales para un grupo y no para otros grupos.

## G) Comentarios finales

Los ítems de la escala Guttman, como los de la escala Lickert, son acumulativos. Sin embargo, los ítems en la escala Lickert son, cada uno de ellos, acumulativos; mientras que en la escala Guttman consiste en una *serie acumulativa de ítems*, donde cada ítem tiene además carácter de acumulativo. De allí que en la escala Lickert sea posible que todos los ítems puedan ocupar aproximadamente el mismo lugar de la escala; en el caso de la escala Guttman, utilizando la técnica H de Edwards y Kilpatrick, garantizamos que los ítems en la escala se distribuyan a lo largo de el continuo de actitud. En términos gráficos:

Figura 8.



Ejemplo de algunos ítems acumulativos que forman una serie acumulativa de ítems. Todos los ítems van en la misma dirección (todos los ítems son de tal tipo que todos los sujetos van con una disposición negativa respondan "SI"). En realidad la inclinación para cada ítem y las distancias entre los ítems varían.

## ¿Variación en el instrumento, en los sujetos, o en ambos?

La escala Guttman ordena tanto a los sujetos como a los estímulos con respecto a un continuo de actitud, es decir que tanto a sujetos como a estímulos se les puede asignar valores de escala. Consecuentemente, la escala Guttman corresponde a un enfoque centrado en la respuesta. En este caso, tanto la actitud del sujeto como la actitud reflejada por el estímulo, actúan para determinar la respuesta del individuo.

MÉTODO DE COMPARACIÓN POR PARES

Consiste en presentar los estímulos al sujeto de a dos por vez (pares) y preguntarle cuál de ellos es el más grande, el mejor, el más cálido, etc. Por medio de este artificio logramos un orden de rango basado en el número de elecciones recibidas por cada ítem.

Como las comparaciones son sistemáticas (el sujeto compara todo estímulo con cada uno de los otros), el resultado final nos indicará cuán consecuentes es el individuo con sus juicios o evaluaciones. En otras palabras, esto nos da la consistencia de las respuestas o la confiabilidad. En una pauta totalmente consistente, debemos esperar que el estímulo que sea más grande, más agradable, más cálido, etc., para el sujeto tendrá  $n - 1$  elecciones, el segundo estímulo  $n - 2$  elecciones, y así sucesivamente hasta llegar al menos grande, que tendrá 0 elecciones.

Se debe tener cuidado con no presentar sistemáticamente un estímulo en primer o segundo lugar, ya que con ello podríamos favorecer o contrariar la elección. Para evitar esto, se colocan los estímulos aleatoriamente, o se administran formas paralelas.

Este método de comparación por pares requiere de mucho tiempo cuando las comparaciones a realizar son numerosas. El número posible de pares a calcular resulta de la fórmula

$$\frac{n(n-1)}{2}$$

donde  $n$  = cantidad de ítems.

Es decir que si queremos comparar 50 ítems, el número de combinaciones que el sujeto deberá evaluar son: 1 225; si utilizamos 100 ítems, el número de pares a comparar serán 4 950.

A) Ejemplo de construcción de una escala basada en el método de comparación por pares

El método de comparación por pares puede ser usado como un test de Tipo A, o como un test de Tipo B. Analizaremos aquí los resultados de un test de Tipo B en el caso de los jueces y como un test de Tipo A en la aplicación a un sujeto.

1) Versión de los jueces. Test de Tipo B

Supongase que queremos determinar el grado relativo de izquierdismo-derechismo de 5 partidos políticos en un país cualquiera, para después poder ordenar el grado de izquierdismo-derechismo en sujetos según sus simpatías políticas por partidos.

En la versión de los jueces vamos a ubicar a los partidos en un orden de rangos, y además—según la evaluación de los jueces— en una escala intervalar en donde un extremo de la escala representará izquierdismo y el otro derechosismo.

A un grupo de 50 jueces solicitamos que ordenen los 5 partidos A, B, C, D y E dando un valor 0 al partido más a la izquierda, un valor 1 al siguiente, etc., hasta un valor 4 al partido ubicado más a la derecha del espectro político.

Analizaremos los resultados de 5 jueces (se entiende que la evaluación de la escala debe hacerse sobre el total de jueces que se utilicen), para simplificar las operaciones. Los resultados de los ordenamientos de nuestros 5 jueces son los siguientes:

Juez	Ordenación de los partidos				
	A	C	B	D	E
1	4	3	2	1	0
2	3	2	4	0	1
3	2	4	3	1	0
4	4	2	2	3	1
5	3	4	4	2	1

El puntaje máximo posible con 5 jueces para un partido es  $\frac{20}{5} = 4$ . El puntaje mínimo es 0.

Los puntajes para los cinco partidos son entonces:

$$A = (4 + 3 + 2 + 4 + 3)/5 = 3.2$$

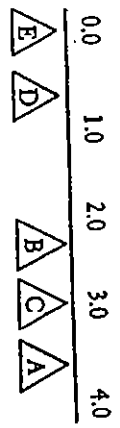
$$C = 15/5 = 3.0$$

$$B = 14/5 = 2.8$$

$$D = 4/5 = 0.8$$

$$E = 1/5 = 0.2$$

El orden de rango de los partidos, de derecha a izquierda es: A - C - B - D - E, y su ubicación en una escala intervalar:



2) Versión de la escala para sujetos. Versión final. Test de Tipo A

En el cuestionario se presentan los partidos políticos por pares, en una lista

en la que se combinan de dos en dos todos los partidos. En nuestro caso los pares a comparar serían:

A-B	B-D
A-C	B-E
A-D	C-D
A-E	C-E
B-C	D-E

Cuidamos por supuesto que los partidos aparezcan "mezclados" en la lista, evitando por ejemplo que el partido A esté siempre al comienzo, y se le sollicita al sujeto seleccione entre cada par a cuál de los dos partidos prefiere.

B) *Ejemplo de una escala de comparaciones por pares en un cuestionario (Entrevista)*

El entrevistador muestra una tarjeta (Hoja II) separada con el contenido arribado con la línea ----- arriba. En el cuestionario aparece el texto total arriba. El entrevistador sigue las instrucciones preguntando cada vez por un solo par.

Las indicaciones se hacen naturalmente en el espacio en el cuestionario y no en la tarjeta (Hoja III). El entrevistador hace la registración en el cuestionario señalando el entrevistado solamente su preferencia.)

Un entrevistado ha contestado de la siguiente manera:

35. Tengo aquí una hoja con los nombres de los partidos políticos, agrupados de dos en dos. Por favor, dígame cuál partido prefiero en cada grupo.	Muestra hoja III
Subtraye el partido escogido, si dos partidos son escogidos al mismo tiempo, subtraye a ambos. Si el entrevistado no puede escoger ningún partido, no subtraye nada.	<p>Conservador - Liberal.          Socialista - Comunista.          Liberal - Partido Laborista.          Laborista - Socialista.          Comunista - Conservador.          Conservador - Socialista.          Liberal - Comunista.          Socialista - Liberal.          Comunista - Conservador.</p> <p>Conserv. 1 = 1          Lib. 2 + 2 + 2 + 2 = 8          Cent. 2 + 2 + 1 = 5          Soc. 2 + 2 = 4          Com. 2 = 2</p>

Las ponderaciones son calculadas después por el entrevistador. Se pone 2 si una de las alternativas fue tomada, y 1 si ambas fueron tomadas. El resultado es el perfil 1, 8, 6, 4, 1 del sujeto. Podemos concluir que la persona simpática más con el partido liberal. La ponderación 2 y 1 que corresponden a cada par es arbitraria. Se puede utilizar, por ejemplo, también 1 y 0. En este caso el perfil sería 0, 4, 3, 2, 0, y el sujeto todavía simpática más con los liberales.

Aparte de ser posible determinar la posición del sujeto en el continuum izquierda-derecha, se puede, por ejemplo, estudiar las personas que tienen inconsistencia en sus perfiles. Además del perfil obtenido podemos calcular la posición del individuo según nuestra escala:

*Resultados para el individuo*

Posición del par- tido a lo largo de la variable	Nombre del partido	Columna: M Núm. de elec- ciones "veridicas"	Posición del indi- viduo a lo largo de la variable
++ 4	A	1	4 por 1 = 4
+ 3	C	8	3 por 8 = 24
2	B	5	2 por 5 = 10
- 1	D	4	1 por 4 = 4
--- 0	E	2	0 por 2 = 0
Total			= 42

El individuo tiene una posición medianamente en el centro de la escala.

Una persona extrema en la variable recibe como máximo 60 puntos y como mínimo 20. Compruebe esto usted mismo a través de cálculos.

Es evidente que es mucho más fácil para el sujeto indicar cuál de las 2 alternativas prefirió más. Sopesar cada ítem en una escala es más difícil como también es complicada la tarea de ordenar por rango los ítems. Cuando el número de ítems es más de 3 ó 4 en la tarea de ordenar por rango, el sujeto mismo tiene que usar alguna especie de comparación de los ítems entre sí. De este modo el método de comparaciones por pares se hace en una forma sistemática. El resultado de este método es un orden de rango basado en el número de elecciones recibidas por cada ítem.

Una ventaja con el método de comparaciones por pares es que muestra cuán "consecuente" es el sujeto en su estimación de las alternativas comparadas. Una inconsistencia se muestra al disminuir irregularmente la serie de números en la columna M. Si el número de ítems es N y si el sujeto es consecuente, nosotros obtendremos (N - 1) elecciones recibidas para el ítem que al sujeto le gusta más, (N - 2) para su segunda preferencia de ítem, etc. Si el número de ítems es 7, así recibimos 6, 5, 4, 3, 2, 1, 0 (en el caso de la persona mencionada anteriormente). 6, 5, 4, 2, 2, 2, es un ejemplo de inconsistencia. La suma de los valores de la serie de números es en ambos casos 21, el número de comparaciones por pares.



Hay algunas variantes del método de comparaciones por pares, como el método de comparaciones por pares dobles y comparaciones triadas que han sido usadas en medición de intereses y actitudes. Estos métodos sin embargo, hacen más difícil que el sujeto haga una ordenación "consecuente" (precisión más baja).

Algunos *test* de interés y actitudes muy conocidos que usan el método de orden por rango o el método de comparaciones por pares son: el *test* de interés de Allport Vernon, Lindzey en "Study of Values" (actitudes hacia la religión, política, arte, ciencia, etc.) y el registro de preferencia Kuder (interés en actividades como deporte, música, etc.).

### C) Ventajas y desventajas

*Ventajas.* a) Proporciona resultados más precisos. b) Muestra claramente cuán "consecuente" es un sujeto con sus propios juicios. c) Es más fácil para un sujeto indicar a cuál de dos alternativas prefiere más.

*Desventajas.* a) Cuando los juicios son numerosos, las comparaciones son muy grandes. En el caso de 20 ítems, el número de comparaciones a realizar es de 190. b) Gran consumo de tiempo de entrevista.

## EL DIFERENCIAL SEMÁNTICO. LA ESCALA DE OSGOOD

### A) Antecedentes teóricos

El método es descrito por los autores como un método para medir el significado que tiene un objeto para un individuo.

Osgood supone que existe un espacio semántico de dimensionalidad desconocida y de naturaleza geométrica. El espacio está construido (o constituido) de escalas semánticas. Cada escala consiste de un par de adjetivos que son bipolares. Se supone que esas escalas forman una función lineal que pasa a través del origen. Para estar en condiciones de definir el espacio adecuadamente, es necesario usar una gran cantidad de escalas que son una muestra representativa extralada del universo de escalas. Para diferenciar el significado de un objeto, el individuo hace una elección entre las alternativas dadas. La función de cada elección es localizar el objeto en el espacio semántico. La validez de la localización en este punto en el espacio depende del número y representatividad de las escalas.

De este modo, la diferencia semántica significa la estabilización sucesiva (anclaje) de un objeto hasta un punto en el espacio multidimensional semántico, a través del puntaje de un número de alternativas semánticas dadas precisadas en la forma de escalas. Una diferencia de significado entre 2 objetos es simplemente una función de las diferencias de su ubicación en el mismo espacio, es decir, una función de la distancia multidimensional entre 2 puntos.

El punto en el espacio que da una definición operacional del significado tiene 2 características principales: 1) Dirección desde el origen; 2) Distancia

desde el origen. Esto podría ser explicado como el tipo e intensidad del significado.

La dirección desde el origen depende de cuál de los polos de la escala se elige y la distancia depende de cuán extrema es la posición elegida en la escala.

### 1) Dimensiones en el espacio semántico

Osgood dio gran importancia al muestreo. El diferencial semántico está influido por 3 fuentes de variación: el individuo, las escalas y los objetos. Muchas diferentes modificaciones fueron hechas para asegurar la universalidad de la estructura del factor (*factor-structure*) pero siempre Osgood obtuvo los mismos factores principales en los diferentes análisis y así llegó a la conclusión de que la estructura del factor no dependía de la elección de escalas. El *test* 8 factores de hecho explicó la mayor parte de la varianza total, mientras otras dimensiones sólo explicaban una pequeña parte de ellas.

### Dimensiones

- 1) La evaluación que hace el individuo del objeto o concepto que se está clasificando. Ejemplo de escalas bipolares: regular-irregular; limpio-sucio; bueno-malo; valioso-despreciable.
- 2) La percepción del individuo de la potencia o poder del objeto o concepto. Escalas: grande-pequeño; fuerte-débil; pesado-liviano.
- 3) La percepción del individuo de la actividad del objeto o concepto. Escalas: activo-pasivo; rápido-lento; frío-caliente.

### C) Construcción

El método para el diferencial semántico no es una prueba con ciertos ítems y puntajes de *test* específicos. Debe ser visto como un método para reunir cierto tipo de información (un método que puede ser generalizado), el cual tiene que constituirse por las demandas que presenta cierto problema de investigación. No hay objetos estándar o escalas estándar.

*Selección del objeto (concepto):* "objeto" se usa para determinar qué significa el "estímulo" que da "reacción" (respuesta) en el individuo a través de su indicación en las escalas de adjetivos.

El objeto puede ser verbal; puede consistir de sólo una palabra o de varias palabras. Objetos no-verbales pueden ser diferenciados (cuadros u otros estímulos estéticos).

La elección correcta de un objeto es un problema de muestreo. Esto generalmente significará en la práctica común que el investigador usa su sentido común al seleccionar el objeto. El investigador debería pensar en elegir objetos que se supone darán: 1) Diferencias individuales (para poder estudiar la variación en el material); 2) Que tengan un solo significado (de otra manera hay riesgo de vacilación en la elección); 3) Se supone que todos los individuos lo conozcan bien (de otro modo habrá regresión al medio de la escala).





que, de acuerdo a las evaluaciones de éstos, sean seleccionados finalmente los ítems considerados adecuados para que integren la escala.

En la práctica, lo que han hecho los investigadores es aplicar directamente la escala ya estructurada por Bogardus, consistente en 7 ítems. O, por el contrario, solamente han adaptado tales ítems —en sus formulaciones— a las necesidades del estudio.

B) *A continuación presentamos los ítems que ilustran esta escala de distancia social*

ESCALA DE DISTANCIA RACIAL

Instrucciones

- 1) Dé su primera reacción ante cada pregunta sin pensarlo demasiado.
- 2) Dé sus reacciones a cada raza considerada como un grupo. No dé sus reacciones de sentimientos a los mejores o a los peores miembros que usted ha conocido, sino piense de acuerdo a la idea que usted tiene de la raza considerada como un todo.
- 3) Ponga una cruz debajo de cada raza en tantas de las siete filas como sus sentimientos lo dicten.

Categoría (ítem)	Judio	Negro	Indio	Blanco europeo	Chino	Mexicano
1. Se casaría con ellos regulares	.....	.....	.....	.....	.....	.....
2. Tendría como amigos regulares	.....	.....	.....	.....	.....	.....
3. Trabajaría en una oficina junto a	.....	.....	.....	.....	.....	.....
4. Viviría en el mismo vecindario con	.....	.....	.....	.....	.....	.....
5. Estaría simplemente hablando como a conocidos a	.....	.....	.....	.....	.....	.....
6. Excluiría de mi vecindario a	.....	.....	.....	.....	.....	.....
7. Excluiría de mi país a	.....	.....	.....	.....	.....	.....

El sujeto es, pues, interrogado para dar sus primeras reacciones de sentimientos y no para racionalizar. La suposición es que las primeras reacciones de sentimientos reflejan actitudes mejor que ningún otro, aparte de la conducta en sí en un período.

Aunque la conducta de larga duración es la mejor prueba de las actitudes de una persona, la escala de distancia social se ideó para dar predicciones, mientras se espera que la conducta de larga duración revele las actitudes. La conducta de un corto período puede revelar pseudactitudes, no acti-

tudes verdaderas. Ella puede medir actitudes que son "ocultas" para ciertos fines.

Se supone, en consecuencia, que las primeras reacciones de sentimientos sin racionalización son significativas en revelar cómo actuaría una persona si ella tuviera que enfrentarse repentinamente con las situaciones citadas en la escala.

C) *Flexibilidad de la técnica*

Una indicación de la flexibilidad de esta técnica puede verse en el hecho de que ella puede ser usada no solamente para escalar grupos o valores que son externos al sujeto que hace el rating, sino también para escalar a los mismos raters con respecto a la distancia social entre ellos y algún grupo. Para ello se utiliza un *continuum* en el sentido de favorabilidad-desfavorabilidad de la escala tipo Likert. Con un ejemplo se ilustrará mejor:

INSTRUCCIONES

- 1) Hablando de los norteamericanos, por favor, ¿podría indicarnos su simpatía o no hacia ellos?
- 2) Por favor, evalúe a los norteamericanos en esta escala, marcando con una cruz en los espacios puntuados la afirmación que expresa su sentimiento hacia ellos.

No los tengo ni simpatía ni antipatía	.....	Me son decididamente simpáticos	.....	Me son un poco simpáticos	.....	Decididamente me son antipáticos	.....	Les tengo un poco de antipatía	.....
---------------------------------------	-------	---------------------------------	-------	---------------------------	-------	----------------------------------	-------	--------------------------------	-------

Esta forma de escala es recomendable y se presenta primero a los sujetos a quienes se les está sometiendo a investigación acerca de sus actitudes. Luego viene la forma anteriormente diseñada de la escala.

La forma ilustrada anteriormente se construye de la misma manera para los demás grupos étnicos, o nacionales, hacia quienes deseamos conocer las actitudes de los sujetos.

D) *Confiability*

Una escala de distancia social no es fácilmente probada para la determinación de su confiabilidad, ya sea por la forma múltiple o por la técnica de la división por mitades (*splithalf*). El enfoque *test-retest* es la medida más efectiva de confiabilidad de tal escala.

E) *Validez*

Para mostrar la validez de la escala se requiere de pensamiento cuidadoso. La aplicación del método del *known-group* implicaría hallar grupos conocidos que sean favorables hacia algunos de los tipos étnicos y no favorables hacia otros. Si las respuestas de estos grupos forman el requisito patrón, entonces la validez parecería probable.

Por otro lado, el uso del método de criterios independientes requeriría que el orden de rango forme algún otro rango de aceptabilidad social. Tales indicadores podrían ser el orden de rango de deseabilidad como inquilinos en un gran proyecto de construcciones de viviendas, de aceptabilidad como miembros de un gremio, etc.

#### F) Limitaciones y aplicaciones

Esta técnica de *scaling* no está limitada en cuanto a su flexibilidad de aplicación ni por su crudeza como medida. Los principales problemas son: 1) La suposición de la equidistancia entre los puntos de la escala; 2) La suposición de que cada punto está necesariamente "más allá" del punto anterior; 3) El hecho de que ella puede ser probada por confiabilidad solamente por el "dejar" de la técnica ordinaria del *test-retest*. Por lo tanto, el uso de este método de *scaling* está limitado generalmente a estudios pilotos o a investigaciones que por alguna razón deben ser completados rápidamente y no requieren de un nivel de precisión muy alto.

#### BIBLIOGRAFÍA RECOMENDADA PARA ESCALAS DE MEDICIÓN DE ACTITUDES

Al lector interesado en profundizar tanto en la teoría como en la técnica de escalas de medición de actitudes recomendamos especialmente los siguientes textos:

- Edwards, A. *Techniques of Attitude Scale Construction*; Appleton-Century-Crofts, Nueva York, 1957.
- Torgerson, W. *Theory and Methods of Scaling*; J. Wiley, Nueva York, 1958.
- Thurstone, L. *The Measurement of Attitudes*, The University of Chicago, 1929.
- Stouffer, S. et al. *Measurement and Prediction. Studies in Social Psychology in World War II*; Princeton University, 1950.
- Upshaw, H. "Actitudes Measurement" en Blalock, H. y Blalock, A. *Methodology in Social Research*; McGraw-Hill, Nueva York, 1968.
- Selluz, Jahoda et al. *Research Methods in Social Relations*; Holt-Rinehart-Winston, Nueva York, 1959.

Recomendamos también revisar especialmente las siguientes revistas:

- Sociology and Social Research*  
*Journal of Applied Psychology*  
*Journal of Social Psychology*  
*Journal of Abnormal and Social Psychology*  
*American Sociological Review*  
*American Journal of Sociology*  
*Psychological Bulletin*

## VII. EL TRABAJO DE CAMPO

INGVAR ABERMAN

Suponemos que los entrevistadores han sido seleccionados y se ha verificado su capacidad para el trabajo en el terreno en el cual va a ser realizado. Supongamos también que el investigador cuenta con una central para el trabajo de terreno, desde la cual él opera. Esta central tiene servicio telefónico, una gran sala de conferencias y algunas salas pequeñas, a las cuales los entrevistadores pueden llevar a los R,<sup>1</sup> si es que existen dificultades para obtener un lugar adecuado para realizar la entrevista en otra parte. En la central existen además planos de la ciudad, guías telefónicas, mapas de comunicaciones, etc.

Algunos días antes de que comience el trabajo en terreno, los entrevistadores deben tener una reunión preliminar, con el objeto de tomar todas las instrucciones necesarias, ejercitarse en algunos cuestionarios, formularios, etc. Algunos de los puntos de importancia que deben tratarse en estas reuniones preliminares, son los siguientes:

### 1) Descripción del estudio

El investigador ofrecerá una descripción metodológica del estudio y explicará algunas de las partes teóricas que hay dentro de él.

### 2) Descripción de la muestra

El investigador dará una explicación acerca de la forma como se ha diseñado la muestra, su tipo y qué posibilidades existen para reemplazar los individuos que no se encuentran al eleccionar las entrevistas.

Los puntos 1 y 2 son meramente informativos para ubicar a los entrevistadores en el tipo de estudio y si el investigador lo considera conveniente, pueden eliminarse. Los puntos que vienen a continuación son bien específicos y conviene se los tenga en cuenta.

### 3) Como ponerse en contacto con los respondientes

A cada entrevistador debe asignársele un distrito especial, propio. El entrevistador recibe un mapa y se le pide que haga un plan para sus contactos, que no ofrezca mucha pérdida de tiempo en movilización. De acuerdo con la profesión del R, el entrevistador selecciona el tiempo que considera más probable para contactarse con las personas elegidas. Es importante que se contacte con el mayor número de R posibles. No siempre existe la seguridad de que la persona disponga del tiempo necesario para la entrevista cuando es visitado por el entrevistador, aunque esté realmente interesado. En este

<sup>1</sup> R = Respondiente.