

Cuadernos de Investigación Social
Colección dirigida por Manuel Mora y Araujo

Wainerman

Catalina H. Wainerman

S. S. Stevens

Robert L. Thorndike

L. Cronbach y P. Meehl

Rensis Likert

L. L. Thurstone

Louis Guttman

Charles E. Osgood

G. J. Suci

P. H. Tannenbaum

*Escalas de medición
en ciencias sociales*

Selección e introducción de Catalina H. Wainerman

Ediciones Nueva Visión

Buenos Aires

Catalina H. Wainerman

Introducción

© 1976 por Ediciones Nueva Visión S.A.I.C.
Tucumán 3748, Buenos Aires, Argentina
Queda hecho el depósito que marca la ley 11.723
Impreso en la Argentina/Printed in Argentina
Prohibida su reproducción total o parcial

El propósito de esta antología es presentar un conjunto de trabajos ya "clásicos" en el área de la medición de actitudes mediante escalas.¹ La selección obedeció a varios criterios. El primero y principal, poner al lector de habla castellana en contacto directo con un conjunto de trabajos originales referidos a la elaboración y aplicación de escalas de medición de actitudes —aquellos cuya frecuencia de mención en la literatura especializada constituye un testimonio de su significación—. El segundo, presentar los modelos de escalas más frecuentemente utilizados en sociología y psicología social —tipo Likert, tipo Thurstone, escalograma de Guttman y diferencial semántico de Osgood, Suci y Tannenbaum—. Finalmente, el tercero, incluir trabajos que no se limitan a exponer los modelos teóricos subyacentes a los cuatro tipos elegidos sino que, además, incursionan en aspectos de procedimiento, los más directamente vinculados a la elaboración y aplicación de los instrumentos.

Cuando un investigador se propone construir una escala, centra su atención sobre un conjunto de hechos observables, especifica la (o las) variable(s) en términos de la cual habrá de describir los elementos del mencionado conjunto y formula las reglas que le permitirán distinguir dichos elementos según la posición que ocupen a lo largo de la variable en cuestión. Cumplido lo anterior, el investigador está en condiciones de representar mediante números las varias posiciones distinguidas por sus reglas. En ese momento puede decirse que la variable —tal como la manifiesta el conjunto de hechos observables— ha sido medida mediante la escala.

Lo dicho permite describir la escala como un instrumento de medición formado por tres tipos de elementos: estímulos, o más bien, aspectos de estímulos (personas, acontecimientos, conceptos, etcétera); números, y reglas que vinculan los estímulos con

¹ Otro volumen de esta misma colección incluye una selección de artículos sobre instrumentos de medición en general en ciencias sociales; se trata de M. Mora y Araujo (comp.), *Medición y construcción de índices*, Nueva Visión, Buenos Aires, 1971.

los números. La aplicación de una escala constituye una instancia de medición —es decir, de asignación sistemática de números a términos de una o más variables. La medición es posible debido a que algunas de las propiedades del sistema de los números —en algún aspecto diferenciable de un conjunto de observaciones—, destinada a reflejar la posición de cada miembro del conjunto en particular, igualdad, orden, igualdad de diferencias y de razones—son isomórficas con algunas de las propiedades de aspectos de los estímulos. A causa de dicho isomorfismo es posible formular reglas que pongan en correspondencia sistemática elementos de los dos sistemas, el formal y el empírico. A cada tipo de reglas y de operaciones empíricas que sea legítimo realizar con los elementos del mundo empírico corresponden diferentes niveles de medición —nominal, ordinal, de intervalo y de razón—.

La temática de la medición y de sus niveles constituye el foco de interés del capítulo de Stevens que abre esta antología. Junto con los trabajos de Russell (1938)² y Campbell (1938),³ el de Stevens representa una de las piezas teóricas fundamentales de la literatura sobre medición en ciencias sociales. El capítulo se inicia con una exposición de los fundamentos matemáticos de la medición y culmina con la descripción de los cuatro tipos de escala (niveles de medición) antes mencionados, en términos de: las operaciones empíricas que admite cada tipo; la estructura matemática del grupo al que pertenecen dichas operaciones y la clase de transformaciones que son posibles (aquellas que dejan invariante la estructura de la escala); y las medidas estadísticas que pueden aplicarse con legitimidad a los datos de la escala (aquellas que preservan su constancia bajo las transformaciones permitidas por cada tipo de escala).

Una vez diseñada y aplicada una escala a un grupo de sujetos, el investigador se encuentra con un conjunto de puntajes. Como es el caso con cualquier instrumento de medición, los puntajes están determinado por —y por lo tanto reflejan— no solo la(s) variable(s) que se pretende medir sino también una variedad de otros factores más o menos estables o más o menos tran-

sitorios que dependen de los sujetos, del instrumento y de su aplicación. Dichos factores que, como ya se dijo, no son exclusivos de la medición con escalas sino compartidos por cualquier tipo de instrumentos de medición en ciencias sociales, tienen efecto sobre dos aspectos esenciales de una buena medición: su confiabilidad y su validez. Un procedimiento de medición es confiable en la medida en que aplicaciones independientes del mismo produzcan resultados consistentes, y es válido en la medida en que los puntajes que origine reflejen “verdaderas” diferencias entre los individuos o grupos a lo largo de la variable (o variables) medida(s), antes que diferencias debidas a otras variables posibles.

La confiabilidad es una condición necesaria, si bien no suficiente, del valor de los resultados de una investigación y de su interpretación. Un alto grado de confiabilidad no es garantía de buenos resultados científicos, pero éstos no son posibles en caso de no ser confiables. Este problema está tratado exhaustivamente en el artículo de Thorndike incluido en esta antología. El trabajo excede largamente la problemática de la medición mediante escalas. Incluye una extensa lista de factores que pueden producir variaciones en un conjunto de puntajes, además de los que producen diferencias sistemáticas respecto de la variable medida. Se trata de factores que atentan contra la confiabilidad del instrumento y que tienen origen, como se dijo, en los mismos sujetos evaluados, en el instrumento de evaluación y en su aplicación. En el capítulo de Thorndike ha de encontrarse, además, una exposición de los tipos de procedimientos disponibles para estimar la confiabilidad de un instrumento —uso de formas equivalentes, administración repetida, división de los ítems de un instrumento en dos grupos y análisis de la variancia entre ítems individuales— y un análisis de sus respectivas ventajas y desventajas.

Dijimos que la confiabilidad de un instrumento es una condición necesaria pero no suficiente de la calidad de los datos producidos mediante su aplicación. Los datos sobre la confiabilidad de un instrumento proporcionan información acerca de la precisión con que éste mide, pero no acerca de si mide lo que se desea. Esta última información —que hace a la validez del instrumento— es básica en cualquier instancia de medición; de ahí la afirmación de Thorndike acerca de que la información sobre la confiabilidad

² B. Russell, *Principles of Mathematics*, Norton, Nueva York, 1938.

³ N. R. Campbell, “Symposium: Measurement and its Importance for Philosophy”, *Proc. Arist. Soc. Suppl.*, 17, 1938, pp. 121-142.

de un instrumento *complementa* la evaluación inicial sobre su validez, lo que significa que sólo tiene sentido hablar de la precisión de una medición una vez determinado su contenido mismo.

El tema de la validez es harto complejo, controvertido y particularmente importante. Su tratamiento requiere la referencia simultánea al nivel teórico y al nivel empírico. Lograr que una medición sea confiable es un problema técnico; asegurar su validez, es más que un problema técnico: hace a la esencia misma de la ciencia —especialmente cuando se trata de la validez conceptual—.

La más importante clasificación de los tipos de validez propuesta hasta el momento fue obra de un comité conjunto organizado por la Asociación Psicológica Norteamericana en colaboración con asociaciones dedicadas a la investigación en el área de la educación.⁴ Dicha clasificación ha servido de base para el trabajo de Cronbach y Meehl incluido en este volumen. En él se presentan cuatro tipos de validez: predictiva, de contenido, concurrente y conceptual, aunque el trabajo está dirigido casi exclusivamente al último.

La validez conceptual representa uno de los avances científicos más significativos de la moderna teoría de la medición. Cuando se pregunta acerca de la validez conceptual de un instrumento, lo que se busca es *explicar* las diferencias entre los puntajes producidos por un instrumento de medición. El interés recae más sobre la variable medida que sobre el instrumento mismo. Lo que diferencia a éste de los otros tipos de validez es la preocupación por la interrelación entre la teoría y los datos empíricos; se trata de un enfoque que contrasta agudamente con los enfoques empíricos que definen la validez de un instrumento estrictamente sobre la base de su éxito en la *predicción* de un criterio.

Los cuatro capítulos siguientes de esta antología presentan —por primera vez en castellano— trabajos originales de los autores de las escalas habitualmente conocidas como “tipo Likert”, “tipo Thurstone”, “escalograma” de Guttman y “diferencial semántico” de Osgood y colaboradores. Las tres primeras pertenecen al conjunto de las escalas de medición de una única dimensión; la

⁴ El trabajo de dicho comité cristalizó en la publicación denominada *Technical Recommendations for Psychological Tests and Diagnostic Techniques*, American Psychological Association, Washington, 1954.

última, al de las escalas multidimensionales. Con excepción de la de Thorndike —que origina una medición de nivel intervalo—, las restantes dan lugar a una medición de nivel ordinal. El diferencial semántico no sólo difiere de las restantes porque posibilita mediciones multidimensionales, sino también porque su elaboración conjuga procedimientos típicos de la construcción de escalas con otros típicos de los tests de asociación. Aun cuando inicialmente se desarrolló para medir el significado (aspectos semánticos) de un estímulo para un individuo, ha llegado a formar parte del conjunto de escalas más frecuentemente utilizado en la actualidad para la medición de actitudes.

La antología se cierra con un capítulo de la autora dedicado a exponer detalles de la aplicación del modelo de la escala de Guttman a la elaboración de un instrumento para la medición de la variable “disposición al cambio”. Dicha escala se utilizó en una investigación sobre la incorporación de innovaciones y el proceso de industrialización argentino realizado entre un grupo de dirigentes de empresas industriales. No se trata de una escala “perfecta”, dista algo de serlo. Su inclusión en este volumen obedeció al deseo de presentar en detalle aspectos de procedimiento relativos a la elaboración de una escala “tipo Guttman” y a otros de relevancia para cualquier tipo de escala.

Catalina H. Wainerman

*Una escala tipo Guttman
de "disposición al cambio"*

Introducción

El objetivo del presente capítulo es familiarizar al lector con los aspectos prácticos de la construcción de una escala tipo Guttman. Más específicamente, se trata de poner al lector en contacto con los aspectos prácticos de 1) la tarea de investigar si los ítems de un conjunto dado son o no indicadores acumulativos de una única variable latente —es decir, si forman una escala unidimensional acumulativa— y 2) la tarea de ordenar a un conjunto de individuos en términos de las posiciones que ocupan a lo largo de dicha variable latente —es decir, adjudicarles puntajes.

Si bien la lógica del método de Guttman no ofrece mayores dificultades de comprensión, a partir de su conocimiento no se infieren directamente los procedimientos de manipuleo de los datos que requiere. Dicho manipuleo, que en el método clásico es bastante complejo, ha logrado simplificarse considerablemente en los últimos años gracias al desarrollo de programas estandarizados de computación (Noland, 1945; Ford, 1950; Toby y Toby, 1954). Dado que dichos programas no se han generalizado aún en nuestro país, expondremos aquí un procedimiento manual que demanda menos manipulación que el método clásico y que tiene la ventaja de ser accesible a todo investigador que no disponga de más recursos que lápiz y papel y cierta dosis de paciencia. La intención primaria de este trabajo es exponer el uso del método de Guttman, cuya utilización en nuestro medio ha sido muy escasa hasta el presente. La intención secundaria es ilustrar sobre algunos aspectos de la elaboración de escalas que son comunes a los varios métodos expuestos en este libro.

Aquí se describirá el proceso de elaboración de una escala que (con pequeñas modificaciones) constituyó uno de los instrumentos centrales en una investigación (Sautu y Wainerman, 1971) dirigida a estudiar las orientaciones sustentadas por un grupo de dirigentes de empresas industriales argentinas frente al cambio tecnológico y la incorporación de innovaciones en sus empresas. El objetivo de la escala en cuestión fue medir el grado de "disposición al cambio" puesta en juego por los entrevistados en el desempeño del rol de empresario. En el contexto específico del estudio, la "disposición al cambio" se definió como una orientación más o menos favorable a considerar, probar y/o adoptar

ideas, técnicas y/o soluciones novedosas en la conducción de una empresa.

El estudio se realizó mediante la técnica de encuesta en un grupo de 107 empresarios seleccionados, de manera no aleatoria, entre dirigentes pertenecientes a seis ramas de actividad industrial. Dichas ramas difieren en términos del período histórico en que comenzaron a operar en la Argentina, y en términos de la tecnología que utilizan —de base predominantemente científica o predominantemente empírica—. Se trata de: hilanderías de lana y de algodón —ambas de origen antiguo y de tecnología predominantemente empírica—, armado de aparatos eléctricos y moldeado de artículos de plástico —ambas de origen reciente y tecnología predominantemente científica— y relaminación de hierro y de acero —ambas de origen intermedio y tecnología mixta—.

1

Elaboración de los ítems de la escala

Como punto de partida, se consideró que la existencia de una disposición favorable al cambio —la dimensión psicológica que se deseaba medir— implica la percepción de su necesidad, la motivación para lograrlo y la preparación para desarrollar comportamientos dirigidos a ese fin. Para conocer en qué medida un grupo tiene una disposición más o menos favorable hacia el cambio, es deseable tener acceso a la observación directa de su conducta de modo de poder inferir a partir de ella dicha disposición. El empleo de tal técnica demanda observar —a lo largo de cierto período de tiempo— el comportamiento de los individuos frente a una variada muestra de situaciones que sean ocasión para la manifestación de tal disposición. Las dificultades prácticas inherentes a este procedimiento son, a todas luces, evidentes. Debido a ellas, en el estudio en cuestión se optó por una técnica sustituta basada en la utilización de tests situaciona-

les. Esto permitió observar las reacciones de los empresarios entrevistados “como si” estuvieran enfrentando situaciones reales.¹

Una vez definida el área psicológica a medir —la disposición al cambio— y elegida la forma de los ítems —tests situacionales— debió seleccionarse el contenido de los mismos —tarea para la que existen pocas normas objetivas puesto que, perteneciendo al contexto del descubrimiento, depende en gran medida de la imaginación del investigador—.

La disposición al cambio puede manifestarse (y por lo tanto medirse) en una multiplicidad de situaciones y frente a una multiplicidad de estímulos. ¿Cuáles elegir? En el caso que nos ocupa se decidió recurrir a un tipo de situación de alta frecuencia en la actividad empresarial: la toma de decisiones. Si bien toda actividad humana lo demanda, la toma de decisiones (en condiciones de mayor o menor incertidumbre) ocupa un lugar central en la actividad empresarial. Por este motivo se decidió que los tests situacionales habrían de plantear situaciones dilemáticas que demandaran tomar una decisión entre un par de alternativas. Estas habrían de implicar diferentes niveles de riesgo y sus resultados, diferentes niveles de cambio.

Una vez tomada esta decisión fue necesario seleccionar (diseñar) un conjunto de ítems del universo de los que potencialmente satisfacen el modelo descripto. Para ello se tomaron en cuenta criterios sustantivos y criterios metodológicos. Desde el punto de vista sustantivo, se hizo explícita la necesidad de medir la disposición al cambio en el mayor número posible de aspectos del rol de empresario de mayor vinculación con la introducción de innovaciones en las empresas. Se trata de aspectos relacionados con las esferas de producción, administración, planeamiento, reclutamiento de personal, etcétera. Desde el punto de vista metodológico, se intentó que los tests satisficieran no solo los criterios comunes a toda escala —relevancia, formulación no

¹ No es éste el lugar para discutir el grado de validez de este “como si”. Se trata de un problema persistente en ciencias sociales que no abordaremos aquí. Solo deseamos destacar que, a nuestro juicio, el uso de tests situacionales aplicados a la medición de actitudes es un instrumento de mayor validez externa que el que se basa en el pedido de la expresión de acuerdo o desacuerdo con un conjunto de enunciados verbales, la técnica más usada en estos casos.

ambigua, capacidad de medir grados variados de la dimensión, etcétera—, sino también que poseyeran un elevado nivel de realismo y de generalidad como para interesar a empresarios ocupados en actividades industriales tan variadas como las incluidas en el estudio.

Para satisfacer los criterios enunciados, en el proceso de elaboración se recurrió a frecuentes consultas con tecnólogos especialistas en las distintas ramas de actividad, a conversaciones informales y a extensos pretests con empresarios ajenos a la muestra estudiada. Esta tarea, siempre la más laboriosa de todo el proceso de construcción de una escala, dio por resultado un conjunto de nueve tests situacionales.

El modelo que subyace a todos los tests es idéntico:² frente a un problema, disyuntiva o propuesta nueva, el individuo debe tomar una decisión entre dos cursos de acción, uno de los cuales supone un nivel mayor de cambio. En todos los casos se describió a los entrevistados la situación, y se los proveyó de un cierto monto de información que en modo alguno era completo. (Estas son las condiciones bajo las cuales habitualmente los empresarios deben tomar decisiones.)

Cuatro de los nueve tests proponen una decisión entre dos cursos de acción cuyos resultados son relativamente anticipables, es decir, presentan un nivel de incertidumbre bajo. La diferencia básica entre ambas alternativas —ambas poco riesgosas— es que una supone cambio y la otra no. Los restantes cinco tests también proponen una elección entre dos cursos de acción, pero, en este caso, ambos implican cambio y, además, uno de ellos es más riesgoso —y objetivamente más atractivo— que el otro. Es decir, los resultados de ambas alternativas suponen introducir cambios pero los de una son menos anticipables y, simultáneamente, en caso de éxito, pueden originar beneficios más interesantes para la empresa.

En resumen, el modelo subyacente a ambos conjuntos de tests consiste en el planteo de una elección entre dos cursos de acción frente a situaciones nuevas. Ambos conjuntos difieren, sin embargo, en cuanto al nivel de riesgo involucrado. En los cuatro

² El modelo de los tests se inspiró en el que subyace a los desarrollados por Kogan y Wallach (1964) para la medición de la disposición a la asunción de riesgos.

primeros tests, el riesgo se acerca a cero; en los últimos cinco, difiere de cero. En el primer caso, aceptar una de las alternativas significa adoptar un comportamiento de cambio y rechazarla, lo contrario. En el segundo caso, elegir una de las alternativas supone un comportamiento de cambio más arriesgado, lo que en el contexto de las situaciones descriptas, supone también mayor nivel de cambio.³

Responder a cada uno de los nueve tests implicaba seleccionar una de cuatro opciones prefijadas. En el caso de los primeros cuatro tests dichas opciones se referían a la alternativa que suponía un mayor nivel de cambio, y variaban en contenido —positivo a negativo— e intensidad —decididamente a probablemente—. (El entrevistado debía indicar cuán dispuesto estaría a adoptar la alternativa que suponía mayor nivel de cambio.) En el caso de los últimos cinco tests, las cuatro opciones se referían a la alternativa objetivamente más atractiva —la que suponía un mayor nivel de riesgo— y variaban en términos de sus probabilidades de éxito. (El entrevistado debía indicar qué nivel de riesgos estaría dispuesto a correr en aras de obtener los beneficios de la alternativa más riesgosa.) Las opciones ofrecidas —expresadas de manera verbal, no numérica— corresponden aproximadamente a los niveles de riesgo 0; 0,10; 0,50 y 0,90, es decir, bajo, moderado y alto.⁴

³ En los nueve casos, luego de haber obtenido las respuestas, se preguntó a los entrevistados por las razones de su elección. Este procedimiento brindó un material de singular riqueza para conocer las bases sobre las cuales los empresarios tomaban sus decisiones y, de ese modo, explorar si el significado atribuido por ellos a las situaciones coincidían con el pretendido. El procedimiento permitió, además, conocer cuáles de los ítems de información incluidos en cada test tenían mayor saliencia para cada entrevistado, y qué monto de discriminación cognitiva y de afecto habían puesto en juego en la elección de sus respuestas.

⁴ Hubiera sido deseable utilizar una escala de riesgos de mayor nivel de discriminación y de intervalos iguales. Esto hubiera requerido presentar las opciones en forma numérica, en términos de probabilidades, en lugar de utilizar la forma verbal que no permite alcanzar tal nivel de discriminación. Esta alternativa se desechó por considerarse que el manejo de probabilidades hubiera presentado mayores dificultades a algunos de los entrevistados de modo que, en lugar de ganar en refinamiento, nos hubiéramos expuesto a ganar en frecuencia de ausencia de respuestas o de respuestas de poca o dudosa validez.

A continuación se transcriben los nueve tests en el orden en que fueron presentados, junto con sus categorías de respuesta.

- 1) "Suponga que usted es el propietario de una pequeña refinería de petróleo. Un día lo visita el representante de una agencia de psicología industrial muy acreditada quien le propone pintar la totalidad de los equipos de su planta asegurándole, en base a su exitosa experiencia, que el cambio aumentará en un 20 % su actual nivel de productividad. El arreglo que le propone es el siguiente: si en el primer mes la productividad aumentara en un volumen equivalente a tres veces el costo de la pintura, la refinería pagaría la totalidad del costo del trabajo, si no aumentara en esa proporción, la refinería pagaría solo un tercio del costo total del trabajo.

¿Estaría usted dispuesto a aceptar el trato?"

rotundamente no _____ probablemente sí _____
lo más probable es que no _____ decididamente sí _____

- 2) "Imagine usted que un ingeniero de reconocida experiencia en su rama ha ideado un nuevo proceso alternativo a uno que usted usa en su empresa. Suponga que este señor le propusiera probarlo en su empresa sin compromiso de adopción para usted. Solamente una vez obtenidos los resultados podría usted considerar la adopción o no del nuevo proceso. Usted no debe temer en modo alguno que el ingeniero utilice el acceso a su empresa para obtener información secreta sobre su funcionamiento. Sabiendo que la prueba demandaría sólo una erogación limitada, ¿le permitiría usted al ingeniero hacer la demostración en su propia empresa?"

rotundamente no _____ probablemente sí _____
lo más probable es que no _____ decididamente sí _____

- 3) "Suponga que en su departamento de producción, que está funcionando satisfactoriamente, usted debe llenar una vacante de nivel intermedio. Se presentan dos personas para el puesto. Una parece muy eficiente, tener muchas ideas y deseos de ponerlas en práctica. La otra parece igualmente eficiente y aparenta tener una alta predisposición para adaptarse rápidamente a la organización de su empresa.

Usted considera que la primera cumplirá eficientemente su trabajo pero que destinará parte de su tiempo a probar sus ideas novedosas que, en caso de éxito, redundarán en beneficio de su empresa. La segunda cumplirá eficientemente su trabajo ciñéndose a lo que se ha hecho hasta el momento en la sección.

¿Elegiría usted a la primera persona para el puesto?"

rotundamente no _____ probablemente sí _____
lo más probable es que no _____ decididamente sí _____

- 4) "Suponga usted que un día lo visita un experto en organización de empresas que pertenece a una firma muy acreditada y conocida por su seriedad y responsabilidad. Usted sabe que bajo ninguna circunstancia haría uso indebido de la información obtenida. El experto ha ideado un nuevo sistema de organización técnico-administrativo que implica el uso de computadoras. Desea probar si ese sistema es practicable en la Argentina para lo cual necesita conocer la organización —no la contabilidad— de varias firmas de plaza.

El experto ha seleccionado su empresa como uno de los casos interesantes a estudiar. Permitírsele no supondría ningún tipo de erogación o de compromiso alguno de adquirir sus servicios, sólo demandaría algo de su tiempo. Una vez recogida suficiente experiencia, y si los resultados fueran ventajosos, la firma a la que pertenece el experto intentará imponer el nuevo sistema en el mercado.

¿Estaría usted dispuesto a permitir que su empresa fuera uno de los casos a estudiar?"

rotundamente no _____ probablemente sí _____
lo más probable es que no _____ decididamente sí _____

- 5) "El señor Gómez es presidente de una empresa manufacturera que está pasando por un período de gran prosperidad y tiene grandes posibilidades de expansión mediante la construcción de una planta en su misma ubicación actual, o en una nueva ubicación en otra región del país. La instalación de la nueva planta en su ubicación actual le supondría un retorno moderado de su inversión. La instalación en la nueva ubicación, en cambio, un retorno mucho mayor debido a menores costos de materia prima, transporte y exenciones impositivas. Por otra parte, la nueva ubicación tiene una larga tradición de huelgas del personal obrero organizado en sindicatos fuertes.

Si usted fuera el señor Gómez, ¿en qué caso se decidiría a instalar la planta en la nueva ubicación?"

en ningún caso _____
solo si juzgara que las posibilidades de huelgas futuras _____
fueran muy escasas _____
cuando juzgara que las posibilidades de que ocurrieran _____
huelgas no fueran mayores del 50 % _____
aun cuando juzgara que las posibilidades de que _____
ocurrieran huelgas fueran mayores del 50 % _____

- 6) "La empresa ABC, una importante firma de plaza, ha contratado por tres años a un químico industrial muy experimentado para su departamento de investigación. La empresa debe decidir cómo emplear los tres años de ese profesional. Podría asignarlo a trabajar en un problema difícil, de largo plazo, que, en caso de poder resolverlo, supondría un adelanto tecnológico muy importante para la firma, o bien optar por

asignarlo a trabajar —como la mayoría de los colegas de la empresa— en una serie de problemas de corto alcance para los que sería más sencillo encontrar soluciones, pero que supondrían adelantos tecnológicos de mucha menor importancia.

Si usted fuera el propietario de la empresa ABC, ¿en qué caso decidiría asignar al químico a la investigación más difícil y de largo plazo?"

- aun cuando las posibilidades de resolver el problema fueran muy escasas _____
- si creyera que las posibilidades de resolver el problema son por lo menos del 50 % _____
- solo si creyera que las posibilidades de resolver el problema son muy altas _____
- en ningún caso _____

7) "El señor Benítez es dueño de una empresa que produce una variada línea de productos y que le asegura un bienestar económico seguro pero modesto. El señor Benítez está considerando la posibilidad de reducir algunas de sus líneas de producción para especializarse en unas pocas de más dinamismo. Esto le permitiría reducir sensiblemente la complejidad de su actual organización y sus costos fijos y, en caso de tener éxito, participar activamente del crecimiento del mercado y aumentar sensiblemente su actual nivel de ganancias. Si fracasara, en cambio, pasaría por un período de serias dificultades para reestructurar su planta y continuar con su actual plan de producción. Si usted fuera el señor Benítez, ¿en qué caso se lanzaría a especializar su planta?"

- en ningún caso _____
- solo si juzgara que las posibilidades de éxito fueran muy altas _____
- cundo juzgara que las posibilidades de éxito fueran del 50 % _____
- no dudaría en hacerlo _____

8) "El señor Aguirre es un industrial de 40 años de edad. Su empresa produce varios artículos. Uno de sus ingenieros acaba de desarrollar un producto totalmente nuevo que no tiene equivalente en plaza. Su fabricación requeriría derivar una parte sustancial de los recursos que la firma dedica actualmente a la producción de uno de sus productos más importantes que le representa del 25 % al 30 % de sus ganancias. Como el producto recientemente desarrollado no tiene equivalente en plaza, el señor Aguirre no tiene forma de estimar cuál será la aceptación en el mercado, pero sí tiene la certeza de que, en caso de ser aceptado, supondría un enorme prestigio para su firma y una ganancia cinco veces superior a la que obtiene con la producción del artículo al que reemplazaría.

Si usted fuera el señor Aguirre, ¿en qué caso se decidiría a encarar la producción del nuevo artículo?"

- aun cuando intuyera que las posibilidades de aceptación por el mercado fueran escasas _____
- cundo intuyera que las posibilidades de aceptación por el mercado fueran tan grandes como las de rechazo _____
- solo si intuyera que las posibilidades de aceptación por el mercado fueran muy altas _____
- en ningún caso _____

9) "El señor Pérez, un ingeniero químico casado y con un hijo está trabajando en una gran firma desde que se recibió, hace cinco años. En la empresa el señor Pérez está haciendo una buena carrera con un sueldo adecuado a su posición.

En ocasión de asistir a una reunión de gente del ramo al señor Pérez le ofrecen empleo en una gran firma recién organizada que, por lo tanto, tiene un futuro incierto. Este nuevo empleo le ofrecería un sueldo inicial igual al que actualmente percibe y además participación en la dirección de la firma, si la empresa sobrevive a la competencia. Si usted fuera el señor Pérez y debiera tomar la decisión, sin duda tomaría en cuenta las posibilidades de éxito de la nueva firma. ¿En qué caso se inclinaría usted a tomar el nuevo puesto?"

- en ningún caso _____
- solo si juzgara que las posibilidades de éxito de la nueva empresa fueran muy altas _____
- cundo juzgara que las posibilidades de éxito de la nueva empresa fueran por lo menos del 50 % _____
- aun cuando juzgara que las posibilidades de éxito de la nueva firma fueran inferiores al 50 % _____

Como puede inferirse a partir de la lectura de los tests, las situaciones que plantean varían a lo largo de una serie de dimensiones. Todas ellas participan de un mismo rasgo: presentan problemas nuevos y son ocasión para desarrollar (o no) conductas innovadoras en la tarea de conducción de una empresa.

2

Recolección de los datos

La serie de nueve tests formó parte de un cuestionario administrado por encuestadores. En la consigna que precedió a la serie se pedía a los entrevistados que trataran de imaginar las situa-

ciones que habían de presentárseles "como si" efectivamente estuvieran enfrentándolas y que indicaran cómo actuarían en cada caso. Se insistió en que interesaba conocer lo que ellos *realmente* harían y no lo que creían *debía* hacerse en cada situación. Se agregó que se debía aceptar la información recibida como veraz y que era necesario que se manejaran exclusivamente con ella (sin solicitar información adicional) aun cuando en ciertas ocasiones la juzgaran insuficiente.

Debido a la longitud de los textos, y para facilitar la tarea de los entrevistados, cada test y sus cuatro opciones se transcribieron en una tarjeta que se entregó para su lectura a los empresarios. Para asegurar que no hubiera omisiones, mientras el entrevistado leía su tarjeta, el entrevistador hacía lo propio en voz alta.

Como ya se dijo (cf. nota al pie 3), luego de obtener la respuesta a cada test el entrevistador preguntaba a su interlocutor por las razones que lo habían guiado a elegir una particular opción de las cuatro que se le habían presentado. Estas respuestas (abiertas) se registraron *verbatim*. En un porcentaje no despreciable de casos (alrededor del 20 % de la totalidad de las respuestas), esta segunda etapa determinó un cambio de las respuestas originales de los entrevistados. En ciertos casos fueron los entrevistados mismos los que, al tener que explicitar razones de su elección, decidieron reconsiderarla. En otros casos se hizo evidente a los investigadores la intención original de los entrevistados de encubrir sus inclinaciones más espontáneas.

Una vez completado este análisis y "limpieza" de las respuestas, se procedió a su codificación. Bajo el supuesto ya mencionado de que en las situaciones en cuestión el mayor nivel de riesgo va junto con un mayor nivel de cambio, se homogeneizaron las opciones a los dos conjuntos de tests, traduciendo las del segundo en las del primero, y asignándoles de manera preliminar un código de "1" a "4", como se indica a continuación.

Tests 5, 6, 7, 8, 9	Tests 1, 2, 3, 4	Código
probabilidad 0	decidido no	1
probabilidad 0,10	probable no	2
probabilidad 0,50	probable sí	3
probabilidad 0,90	decidido sí	4

Cuadro 1

Distribución de las respuestas de la muestra de los nueve ítems

Alternativa de cada test cuyos resultados implican mayor nivel de cambio	Respuestas positivas		Respuestas negativas	
	4	3	2	1
1) Pintar los equipos para aumentar productividad	20 %	46 %	15 %	19 %
<i>n</i> = (107)	(22)	(50)	(15)	(20)
2) Permitir acceso a la planta para probar nuevo proceso	38 %	51 %	5 %	6 %
<i>n</i> = (107)	(44)	(51)	(3)	(9)
3) Llenar una vacante del departamento de producción con una persona de ideas	36 %	29 %	17 %	18 %
<i>n</i> = (107)	(38)	(32)	(18)	(19)
4) Permitir a un experto probar un nuevo sistema técnico-administrativo ..	40 %	35 %	17 %	8 %
<i>n</i> = (107)	(43)	(37)	(18)	(9)
5) Instalar una planta en zona amenazada de huelgas	3 %	31 %	22 %	44 %
<i>n</i> = (107)	(3)	(33)	(24)	(47)
6) Asignar un químico a una investigación de largo plazo	10 %	57 %	19 %	14 %
<i>n</i> = (107)	(11)	(60)	(21)	(15)
7) Especializar la producción de la empresa	27 %	34 %	18 %	21 %
<i>n</i> = (107)	(29)	(36)	(20)	(22)
8) Derivar recursos de la empresa a un nuevo producto	14 %	35 %	24 %	27 %
<i>n</i> = (107)	(15)	(39)	(25)	(28)
9) Aceptar un empleo en una nueva firma	17 %	45 %	13 %	25 %
<i>n</i> = (107)	(18)	(48)	(13)	(28)

En otras palabras, las respuestas a los nueve tests se expresaron de modo que variaran en contenido —aceptación-rechazo— y en intensidad —alta-baja—, es decir, en términos de la intensidad de adhesión (o rechazo) al cambio. Las respuestas codificadas como "4" corresponden a las más intensamente favorables hacia el cambio, y las codificadas como "1", a las más intensamente desfavorables. Como se verá más adelante, el método de Guttman provee de procedimientos para decidir si las respuestas codificadas como "3" y "2" (en nuestro caso "probablemente sí" y "probablemente no", y en otros "indiferentes", "neutro", "indeciso", y formas similares) están más cerca del polo positivo ("4") o del negativo ("1").

El cuadro 1 de página anterior resume la distribución de las frecuencias marginales de las respuestas de los 107 entrevistados a los nueve ítems.

3

Evaluación de la adecuación de los datos al modelo de la escala de Guttman

Una vez obtenidos los datos, son varios los procedimientos disponibles para poner a prueba su adecuación a los supuestos del modelo de Guttman. El procedimiento clásico es el conocido bajo el nombre de "técnica Cornell"⁵ que permite realizar, mediante lápiz y papel, las operaciones posibles con el tablero de escalograma.⁶ La técnica Cornell requiere la construcción de una tabla en la que las columnas representan los ítems y sus categorías de respuesta, y las filas las pautas de respuesta de los sujetos a los ítems. A partir de los datos originales, y mediante arreglos sucesivos, se llega a una tabla definitiva que contiene la combinación de las categorías adyacentes de respuesta y el ordenamiento final

⁵ Esta técnica ha sido descrita en detalle por Guttman (1947). Una adecuada descripción resumida puede verse en Edwards (1957, pp. 178-184).

⁶ Técnica desarrollada por Guttman y descrita en detalle por Suchman (1950, cap. 4).

de los sujetos en términos de sus puntajes. El procedimiento para arribar a la tabla final (sobre la que se calcula la adecuación de la escala mediante el coeficiente de reproducibilidad) es enormemente laborioso. Aparte de un conjunto de criterios, no existen reglas de procedimiento precisas para realizarlo de modo que una parte importante de la tarea queda librada a la subjetividad e ingenio del investigador.

Para superar en parte estas dificultades, en el caso de la escala que aquí se presenta, se utilizó una combinación de dos procedimientos que reducen la complejidad y subjetividad inherentes a la técnica de Cornell. Se trata de 1) la matriz de intercorrelaciones de las respuestas a los ítems y 2) la técnica de Goodenough (1944). La primera se utilizó para hacer una primera evaluación de los ítems, para depurar los inadecuados y para agrupar las categorías de respuesta de modo de mejorar la reproducibilidad. La segunda sirvió para evaluar el coeficiente de reproducibilidad de la escala, y para analizar la distribución de las pautas de respuesta en error o fuera de escala, es decir, para evaluar la adecuación de la escala al modelo teórico.

(En rigor, el uso de la técnica de la matriz de intercorrelaciones en una primera etapa no es indispensable. Podría utilizarse solamente la técnica de Goodenough a lo largo de todo el proceso, es decir, también para agrupar las categorías de respuesta y para depurar los ítems. Lo que se gana al hacer uso de la matriz de intercorrelaciones es reducir la tarea de manipulación de los datos y aumentar la objetividad.)

1) Prueba de la unidimensionalidad de los ítems

El procedimiento de la matriz de intercorrelaciones⁷ se basa en el modelo teórico de la "celda cero". Dicho modelo parte del supuesto de que, si los ítems formaran una escala unidimensional acumulativa perfecta, y si los entrevistados respondieran "correctamente"

⁷ Procedimiento que se atribuye a Stouffer y que ha sido expuesto en detalle por Toby y Toby (1954).

tamente", nadie que diera una respuesta favorable a un ítem de mayor dificultad de aceptación, habría de dar una respuesta desfavorable a un ítem de menor dificultad de aceptación. El grado de "dificultad" de los ítems queda definido, para cada muestra particular, sobre la base de las frecuencias marginales de dichos ítems. (En otras palabras, en la técnica de Guttman; los entrevistados juegan un doble rol simultáneamente, el de sujetos cuya actitud se mide, y el de jueces del grado de facilidad o dificultad de aceptación de los ítems.) Los ítems que reciben menor frecuencia de respuestas favorables a la dimensión medida son los más difíciles (expresan un mayor grado de la actitud medida) y los que reciben la frecuencia mayor, los más fáciles (expresan un menor grado de la actitud medida).

El modelo teórico de la celda cero supone la inexistencia de respuestas en aquella celda determinada por la intersección de la categoría de respuesta positiva al ítem más difícil con la categoría de respuesta negativa al ítem que le sigue en dificultad. En el caso de dos ítems dicotómicos, dicho supuesto puede representarse del modo siguiente:

		ítem 1 (más difícil)	
		+	-
ítem 2 (más fácil)	+		
	-	0	

Dado que habitualmente las respuestas de los sujetos no son "perfectas" —o que la unidimensionalidad de los ítems no lo es— no ha de esperarse encontrar celdas cero perfectas. Guttman ha fijado (arbitrariamente) el error máximo aceptable en 10 %, porcentaje fijado sobre el total de entrevistados. En nuestro caso se trata del 10 % de 107 individuos, lo que hace que la mayor frecuencia admisible en la celda cero sea de 11 casos.

El modelo recién expuesto se aplica a un solo par de ítems. La técnica de la matriz de intercorrelaciones consiste básicamen-

te en la aplicación de dicho modelo a todos los pares de ítems del conjunto cuya unidimensionalidad acumulativa se evalúa. Esto significa analizar la totalidad de las tablas bivariantes que surgen del cruce de cada uno de los ítems con todos los restantes, en nuestro caso 36.

Si los ítems son dicotómicos (como en el modelo expuesto), el procedimiento es relativamente sencillo. Consiste en observar si la frecuencia de casos presente en las celdas que ocupan la posición prevista para la celda cero en todas las tablas bivariantes es inferior al 10 % del total. Pero, en general, el caso es que los ítems proponen tres, cuatro y cinco categorías de respuesta. El mismo razonamiento subyacente al modelo de la celda cero es aplicable en estos casos, pero el procedimiento se torna sustancialmente más complejo. Por este motivo, y aunque esto suponga perder algo de información, generalmente se procede a agrupar las categorías de respuesta de los ítems de modo de reducirlas a dicotomías.

El problema que se introduce entonces es el de dicotomizar los ítems. Y las dicotomizaciones posibles son varias. En caso de que los ítems formen escala, alguna (o algunas) de ellas habrán de revelar la existencia de celdas cero adecuadas. La cuestión reside entonces en cuál de las dicotomizaciones elegir. Esto es lo que trataremos ahora.

El procedimiento consiste en diseñar una matriz que contenga la tabulación cruzada de todos los ítems por todos los ítems. En dicha matriz los ítems se ordenan de izquierda a derecha y de arriba hacia abajo en orden decreciente de dificultad (es decir, en orden creciente de frecuencia de respuesta positiva). Dentro de cada ítem, sus categorías de respuesta se ordenan, sobre la base de su contenido manifiesto, en forma decreciente de intensidad de adhesión. En el caso que nos ocupa, y como puede verse en el cuadro 2, los ítems se ordenaron horizontal y verticalmente a partir del que recibió 34 % de respuestas aparentemente positivas ("decidida" y "probablemente" sí) hasta el que recibió 89 % de respuestas aparentemente positivas, es decir, desde el ítem 5 hasta el ítem 2. Las categorías de respuesta de cada ítem

Las líneas trazadas en la matriz indican la dicotomización a que se arribó para cada ítem luego de ensayar un buen número de diferentes agrupaciones. Varias de ellas revelaron números variables de celdas con frecuencias adecuadas al modelo teórico —frecuencias inferiores a los 11 casos en las celdas ubicadas en el rincón inferior izquierdo de las tablas—. La elección de la dicotomización final de cada ítem, en consecuencia, demanda la utilización de algunos criterios adicionales, además del de la celda cero.

Un primer criterio requiere maximizar el número de ítems cuyos cruces con los restantes revele celdas cero. (Los ítems que no se comportan de acuerdo con el modelo, es decir, que originan frecuencias mayores que las admisibles, no forman escala y, por lo tanto, deben excluirse.)

Un segundo criterio tiene que ver con el tamaño de los marginales. Los ítems con marginales muy extremos —más del 90 % o menos del 10 % de los casos de una misma categoría de respuesta— contienen muy poca información, es decir, contribuyen poco a discriminar grados de la actitud de los miembros de la muestra y por lo tanto deben excluirse.

Finalmente, un tercer criterio se refiere a la amplitud de los marginales, criterio que se relaciona con el poder de discriminación de toda la escala (y no de cada ítem particular, como en el caso anterior). Si la mayoría de los ítems fueran muy extremos, es decir, muy difíciles de aceptar —menos del 10 % de la muestra diera una respuesta favorable— o muy fáciles —más del 90 % de la muestra diera una respuesta favorable— el poder de discriminación de la escala sería prácticamente nulo (y su coeficiente de reproducibilidad artificialmente alto). Es conveniente, en consecuencia, que los marginales de los ítems no sean demasiado extremos, que el conjunto incluya variedad de marginales y, de ser posible, algunos próximos al 50 % de modo de obtener una escala que discrimine suficientes grados de la actitud que intentá medir.

El análisis de las celdas en posición inferior izquierda (celda cero) de todas las tablas bivariadas de la matriz indicó que las frecuencias contenidas en la mayoría de ellas no excede la máxima frecuencia admisible de errores (11 casos). En efecto, de

las 36 tablas bivariadas producto del cruce de los nueve ítems, solo cuatro exceden dicha frecuencia; son las determinadas por el cruce de los ítems 3 y 7 (13 casos), 3 y 9 (14 casos), 3 y 1 (16 casos), y 6 y 1 (15 casos). Si se juzga exclusivamente con el criterio de la celda cero, este resultado es altamente satisfactorio, el monto de error es escaso.

La existencia de más de 11 errores en cuatro celdas indicó que algunos de los ítems que los determinaban no formaban escala con los restantes, es decir, para la particular muestra investigada, no eran indicadores de la misma dimensión. Para purificar la escala dichos ítems debieron eliminarse. El problema que se presentó, entonces, fue cuál elegir de los cinco ítems involucrados en la producción del error —1, 3, 6, 7 y 9—.

Los criterios utilizados ya se mencionaron. Como se ha dicho, el más definitorio es el relativo al número de celdas en error que un ítem determine. En nuestro caso, los más indicados parecieron ser los ítems 3 y 1 puesto que el primero producía celdas en error (conteniendo más de 11 casos) con tres ítems y el segundo con dos ítems, en tanto los restantes lo hacían con un solo ítem.

Otros criterios tomados en cuenta son: el número de ítems a eliminar y las frecuencias marginales de cada ítem susceptible de eliminación. La validez de una escala está positivamente asociada con el número de ítems que contiene, por lo tanto es conveniente que el proceso de eliminación afecte al menor número posible de ítems. En el caso presente, la aplicación de este criterio recomendó nuevamente eliminar el ítem 3 puesto que, de no hacerlo, deberían eliminarse tres ítems, el 1, el 7 y el 9. Para el ítem 1, en cambio —habiéndolo eliminado el 6—, habría alternativas puesto que en su reemplazo podría eliminarse el ítem 6 sin aumentar el número de ítems perdidos en la tarea de purificación de la escala. Y es entonces cuando juega el criterio referido a las frecuencias marginales. Este llevó a eliminar los ítems con frecuencias marginales muy extremas (que, como se dijo, aportan muy poca información) o muy similares a las de otros ítems ya incluidos en la escala (redundantes en términos de su capacidad de discriminación, si bien no en términos de su contenido sustantivo). En nuestro caso, este criterio recomendó eliminar el ítem 6 cuya frecuencia marginal (86 %) es más extrema que la del

ítem 1 (81 %).¹⁰ En resumen, el criterio de la celda cero determinó la eliminación de dos de los nueve ítems, el 3 y el 6.

Aun cuando siete de los nueve ítems probaron formar escala, quedaba por evaluar la bondad de la misma. En efecto, como ya se mencionó, satisfacer el criterio de la celda cero es un requisito necesario pero no suficiente para determinar si un conjunto de ítems constituye una buena medida (escala) de la variable latente que intenta calibrar. Además de un adecuado coeficiente de reproducibilidad —tema que se tratará en 2)— la técnica de Guttman impone requisitos a los marginales de los ítems que hacen a la capacidad de discriminación de la escala. Ya se mencionó la conveniencia de evitar marginales muy extremos y de incluir en cambio una variedad de marginales, algunos de los cuales estén próximos al 50 %. Dichos criterios se aplicaron en la tarea de seleccionar qué ítems convendría excluir. También deben aplicarse a los ítems seleccionados para incluir. Hacerlo requiere un examen de los marginales.

El cuadro 3 resume los siete ítems que —según el criterio de la celda cero— forman escala, junto con la dicotomización final de sus categorías de respuesta y las frecuencias marginales correspondientes. (Vale la pena notar que la combinación de categorías de respuesta es bastante diferente de la obtenida sobre la base del contenido manifiesto, criterio que se utilizó para hacer la clasificación inicial resumida en el cuadro 1.)

En el cuadro 3 los ítems se ordenaron en forma decreciente de dificultad, es decir, en orden creciente de frecuencia de respuestas positivas. El análisis de las frecuencias marginales indicó que el ítem 5 es muy extremo (3 %) y que los ítems 2 y 4 juegan un rol muy similar dentro de la escala (89 % y 92 %). El primero de ellos prácticamente no discrimina y, por lo tanto, se lo eliminó.¹¹ Los dos últimos son redundantes y, por lo

¹⁰ Si bien es cierto que el ítem 6 produce en total menos errores que el 1 (57 y 68 respectivamente), es preferible conservar estos once errores más a cambio del aumento que se logra en el poder de medición de la escala desechando un ítem muy extremo.

¹¹ En el trabajo original, este ítem fue incluido por error. Dado que prácticamente no discrimina, su inclusión no modificó casi en nada los resultados posteriores. Solo elevó artificialmente el coeficiente de reproducibilidad.

Cuadro 3

Distribución de los marginales positivos y negativos de los siete ítems que forman escala

Ítems	Respuestas positivas			Respuestas negativas		
	categorías	n	%	categorías	n	%
5	4	3	3	3-2-1	104	97
8	4	15	14	3-2-1	92	86
7	4-3	65	61	2-1	42	39
9	4-3-2	79	75	1	28	25
1	4-3-2	87	81	1	20	19
2	4-3	95	89	2-1	12	11
4	4-3-2	98	92	1	9	8

tanto, conservar ambos en la escala no hubiera aumentado su poder de discriminación más allá del que se derivaría de la inclusión de uno solo de ellos. Hubiera originado, en cambio, una extensión artificial de la amplitud del puntaje de la escala,¹² y por lo tanto se eliminó uno de ellos. La elección recayó sobre el ítem 4 cuya frecuencia marginal es más extrema.

Resumiendo, con las eliminaciones efectuadas, la elección recayó sobre cinco ítems cuyos marginales (dicotomizados) varían entre 14 % y 89 %. Se trata de un conjunto suficientemente aceptable, en principio, aunque hubiera sido deseable que contuviera uno o dos ítems más con marginales entre 30 % y 50 %. Esto hubiera permitido obtener una medida que discriminara más finamente a la muestra estudiada, especialmente al sector con un grado más alto de "disposición al cambio". Pero, antes de aceptar la escala, fue necesario evaluar con mayor precisión la adecuación del conjunto de los cinco ítems finales. Este es el objetivo de la próxima sección.

¹² Suponiendo que se asignara un puntaje de "1" a cada respuesta positiva, la inclusión de seis ítems originaría una amplitud de puntaje entre "0" y "6"; la inclusión de cinco, una entre "0" y "5". Si se hubiera incluido el último ítem, el puntaje de 92 % de los individuos hubiera aumentado sistemáticamente en uno.

2) Evaluación del coeficiente de reproducibilidad de la escala

El concepto de reproducibilidad es clave en la técnica de Guttman. En una escala perfecta, a partir del conocimiento del puntaje obtenido por cada individuo es posible reproducir sus respuestas a todos los ítems de la misma. El problema de evaluar el grado de adecuación de una escala se convierte entonces en el de la medición del monto de error que contienen los ítems elegidos, es decir, en la medición del apartamiento del instrumento del modelo ideal. El procedimiento consiste, básicamente, en confrontar los datos obtenidos con dicho modelo ideal. De las varias técnicas disponibles para hacerlo, la que se ha utilizado en el caso que nos ocupa es la denominada técnica de Goodenough (1944), cuya manipulación es relativamente simple y cuyo cómputo de los errores es tal que hace del coeficiente de reproducibilidad una medida más adecuada que, por ejemplo, la que produce la técnica Cornell. En lo que sigue se describirá la aplicación de la técnica de Goodenough a los cinco ítems seleccionados y dicotomizados mediante el procedimiento de la celda cero.

La técnica de Goodenough requiere la construcción de una tabla de puntajes en la que, como en el cuadro 4, cada columna corresponde a un ítem y cada fila a un individuo. En otras palabras, se trata de una tabla en la que el comportamiento de cada ítem se lee verticalmente y el de cada individuo horizontalmente. Además de las columnas destinadas a cada ítem, la tabla incluye una para los puntajes de los individuos y otra para el cómputo de los errores.

La tabla se construye ordenando a los ítems en forma decreciente de dificultad y a los individuos en forma decreciente de puntaje. Dicho puntaje es igual a la sumatoria de las respuestas positivas dadas por cada individuo a cada ítem (respuestas a las que se asigna el valor "1", en tanto a las negativas se les asigna el valor "0"). En nuestro caso, el puntaje tiene una amplitud de "0" —quienes respondieron negativamente a los cinco ítems— a "5" —quienes respondieron positivamente a los cinco ítems. Las respuestas de los individuos a los ítems se registraron en términos de "0" y "1". Como control de la tabulación, puede verificarse que la suma de los puntajes obtenidos por todos los individuos iguala

Cuadro 4
Asignación de puntaje mediante la técnica de Goodenough

Sujetos	Ítems					Puntaje	Error	Sujetos	Ítems					Puntaje	Error
	8	7	9	1	2				8	7	9	1	2		
1	1	1	1	1	1	5	0	55	0	1	1	0	1	3	2
2	1	1	1	1	1	5	0	56	0	1	1	0	1	3	2
3	1	1	1	1	1	5	0	57	0	1	1	0	1	3	2
4	1	1	1	1	1	5	0	58	0	1	1	0	1	3	2
5	1	1	1	1	1	5	0	59	0	1	0	1	1	3	2
6	1	1	1	1	1	5	0	60	0	1	0	1	1	3	2
7	1	1	0	1	1	4	2	61	0	1	1	1	0	3	2
8	1	1	0	1	1	4	2	62	0	1	1	1	0	3	2
9	1	1	0	1	1	4	2	63	0	1	1	1	0	3	2
10	1	1	1	1	0	4	2	64	0	1	1	1	0	3	2
11	1	0	1	1	1	4	2	65	0	0	1	1	1	3	0
12	1	0	1	1	1	4	2	66	0	0	1	1	1	3	0
13	1	1	1	0	1	4	2	67	0	0	1	1	1	3	0
14	1	1	1	0	1	4	2	68	0	0	1	1	1	3	0
15	0	1	1	1	1	4	0	69	0	0	1	1	1	3	0
16	0	1	1	1	1	4	0	70	0	0	1	1	1	3	0
17	0	1	1	1	1	4	0	71	0	0	1	1	1	3	0
18	0	1	1	1	1	4	0	72	0	0	1	1	1	3	0
19	0	1	1	1	1	4	0	73	0	0	1	1	1	3	0
20	0	1	1	1	1	4	0	74	0	0	1	1	1	3	0
21	0	1	1	1	1	4	0	75	0	0	1	1	1	3	0
22	0	1	1	1	1	4	0	76	0	0	1	1	1	3	0
23	0	1	1	1	1	4	0	77	0	0	1	1	1	3	0
24	0	1	1	1	1	4	0	78	0	0	1	1	1	3	0
25	0	1	1	1	1	4	0	79	0	0	1	1	1	3	0
26	0	1	1	1	1	4	0	80	0	0	1	1	1	3	0
27	0	1	1	1	1	4	0	81	0	1	0	0	1	2	2
28	0	1	1	1	1	4	0	82	0	0	1	1	0	2	2
29	0	1	1	1	1	4	0	83	0	0	1	0	1	2	2
30	0	1	1	1	1	4	0	84	0	0	1	0	1	2	2
31	0	1	1	1	1	4	0	85	0	0	1	0	1	2	2
32	0	1	1	1	1	4	0	86	1	0	0	1	0	2	2
33	0	1	1	1	1	4	0	87	0	0	0	1	1	2	0
34	0	1	1	1	1	4	0	88	0	0	0	0	1	2	0
35	0	1	1	1	1	4	0	89	0	0	0	1	1	2	0
36	0	1	1	1	1	4	0	90	0	0	0	1	1	2	0
37	0	1	1	1	1	4	0	91	0	0	0	1	1	2	0
38	0	1	1	1	1	4	0	92	0	0	0	1	1	2	0
39	0	1	1	1	1	4	0	93	0	0	0	1	1	2	0
40	0	1	1	1	1	4	0	94	0	0	0	1	1	2	0
41	0	1	1	1	1	4	0	95	0	0	0	1	1	2	0
42	0	1	1	1	1	4	0	96	0	0	0	1	1	2	0
43	0	1	1	1	1	4	0	97	0	0	0	1	1	2	0
44	0	1	1	1	1	4	0	98	0	0	0	1	1	2	0
45	0	1	1	1	1	4	0	99	0	0	0	1	1	2	0
46	0	1	1	1	1	4	0	100	0	1	0	0	0	1	2
47	0	1	1	1	1	4	0	101	0	1	0	0	0	1	2
48	0	1	1	1	1	4	0	102	0	0	0	0	1	1	0
49	0	1	1	1	1	4	0	103	0	0	0	0	1	1	0
50	0	1	1	1	1	4	0	104	0	0	0	0	1	1	0
51	0	1	1	1	1	4	0	105	0	0	0	0	0	0	0
52	0	1	1	1	1	4	0	106	0	0	0	0	0	0	0
53	0	1	1	0	1	3	2	107	0	0	0	0	0	0	0
54	0	1	1	0	1	3	2								
								7	15	65	79	87	95	341	56

la suma de las frecuencias de respuesta positiva ("1") obtenidas por los cinco ítems.

El cuadro 4 contiene los datos necesarios para evaluar el coeficiente de reproducibilidad de los cinco ítems seleccionados y para analizar la forma de la distribución de las pautas de respuesta fuera de escala (en error). Antes de hacerlo, vale la pena recordar que dicho coeficiente es una medida de la correspondencia entre la distribución multivariable obtenida y la distribución multivariable esperada en el caso de que los ítems formaran una escala perfecta. Dicho de otra manera, lo que el coeficiente de reproducibilidad mide es la proporción de respuestas que puede predecirse correctamente a partir de los puntajes recibidos por los individuos en la escala.

Si los ítems investigados formaran una escala acumulativa perfecta, ordenándolos en forma decreciente de dificultad, las pautas de respuesta que corresponderían a cada puntaje, expresadas en forma binaria, serían las siguientes:

<u>Pautas de respuesta</u>	<u>Puntaje</u>
0 0 0 0 0	0
0 0 0 0 1	1
0 0 0 1 1	2
0 0 1 1 1	3
0 1 1 1 1	4
1 1 1 1 1	5

El cómputo de errores en la técnica de Goodenough se basa en la comparación entre las pautas de respuestas esperadas y las pautas de respuesta obtenidas para cada puntaje. Cada desviación de una respuesta observada respecto de la respuesta esperada se computa como un error. Lo que se registra en la columna de "errores" es la suma de los errores de cada individuo. En el cuadro 4 puede observarse, por ejemplo, que la pauta de respuesta observada del individuo 11, que tiene un puntaje "4", es 1 0 1 1 1. En cambio, la pauta de respuesta esperada para un puntaje teórico de "4", es 0 1 1 1 1. El investigador que, sobre la base del puntaje "4" del individuo 11, predijera que respondió favorablemente

a los ítems 2, 1, 9 y 7 y negativamente al ítem 8, cometería dos errores puesto que dicho individuo en realidad respondió favorablemente a los ítems 2, 1, 9 y 8 y desfavorablemente al ítem 7. (Como se ve, se trata de errores referidos a la *ubicación* de las respuestas puesto que siempre el *número* de respuestas positivas observadas y esperadas coinciden.) La pauta de respuesta del individuo 11 contiene, por lo tanto, dos errores, cifra que se registra en la columna de errores, en la fila correspondiente.

El coeficiente de reproducibilidad de una escala es el resultado de la diferencia entre la unidad y la proporción de errores que contiene la escala —proporción que es igual al cociente entre los errores observados y la totalidad de errores posibles (número de ítems por número de individuos)—. En nuestro caso dicho coeficiente resultó igual a:

$$1 - \frac{56}{5 \times 107} = 0,90$$

El valor obtenido de 0,90 está dentro de los límites de la amplitud aceptable para el coeficiente de reproducibilidad de una escala unidimensional, de acuerdo con los criterios de Guttman —0,90 a 1.

Un último punto merece atención. Aun cuando, como en el caso que nos ocupa, el coeficiente de reproducibilidad alcanzado por un conjunto de ítems sea igual o superior a 0,90, y aun cuando los restantes criterios atingentes a las frecuencias marginales se cumplan, debe tomarse en cuenta otro criterio antes de concluir que los ítems analizados constituyen una escala. Se trata de la naturaleza de los errores que contiene la escala.

Los errores que mide el coeficiente de reproducibilidad pueden ser de naturaleza aleatoria o sistemática. En caso de ser aleatorios, se suponen indicadores de múltiples variables de escasa relevancia. En caso de ser sistemáticos, en cambio, se suponen indicadores de alguna otra variable de relevancia, además de la que se supone mide la escala. Es la frecuencia de dichas pautas lo que indica si se trata de uno u otro caso. En el primero de ellos, la frecuencia de ocurrencia de cada tipo de pauta de respuesta fuera de escala es baja; en el segundo, es alta (los errores se concentran en pocos tipos de escala), a pesar de lo cual el

coeficiente de reproducibilidad no llega a ser inferior a los límites aceptables.¹³

Para determinar cuál es la naturaleza —aleatoria o no— de los errores contenidos en una escala, es necesario examinar las pautas en error correspondientes a cada puntaje. En el caso que estamos analizando, el número de pautas fuera de escala suma 30. Su distribución, clasificada según el puntaje, es fácil de describir teniendo en cuenta los datos contenidos en la columna de "error" del cuadro 4. Dicha distribución se resume en el cuadro 5.

Cuadro 5

Frecuencia de tipos de pauta de respuesta en error clasificados por puntaje

Puntaje	Pauta en error					Frecuencia de pauta en error
4	1	1	0	1	1	3
	1	1	1	1	0	1
	1	0	1	1	1	2
	1	1	1	0	1	2
3	0	1	1	0	1	6
	0	1	1	1	0	4
	0	1	0	1	1	2
2	0	1	0	0	1	1
	0	0	1	1	0	1
	0	0	1	0	1	3
	1	0	0	1	0	1
1	0	1	0	0	0	2

Como puede verse en el cuadro 5, la distribución de las pautas de respuesta fuera de escala resultó ser relativamente dispersa como para concluir que los errores eran de naturaleza aleatoria y debidos al efecto de una multiplicidad de variables

¹³ Hay un tercer caso de particular importancia —el de la *quasi* escala, que corresponde al de áreas no escalables—. En ese caso generalmente el coeficiente de reproducibilidad no alcanza el nivel aceptable dado que la frecuencia de error es alta, pero las pautas de respuesta fuera de escala ocurren al azar.

de poca relevancia antes que al de alguna otra variable importante presente en el área.¹⁴ En consecuencia, también este criterio proporcionó evidencias favorables a la aceptabilidad de los cinco ítems.

En resumen, las manipulaciones descritas indicaron la existencia de un conjunto de cinco ítems adecuado a los criterios de Guttman. (Como ya se dijo, hubiera sido deseable disponer de uno o dos ítems adicionales con frecuencias de respuesta positiva entre 30 % y 50 % para lograr una mayor discriminación del sector de empresarios con mayor disposición al cambio.) El conjunto seleccionado contiene cinco ítems dicotómicos cuyos marginales varían entre un 14 % y un 89 %; su coeficiente de reproducibilidad alcanza a 0,90 y tiene una distribución aleatoria de las pautas de respuesta en error. La amplitud del puntaje de la escala varía entre "0" y "5". El puntaje asignado a cada individuo —como se indicó en la columna de puntaje del cuadro 4— es igual a la frecuencia de respuestas positivas dadas a los cinco ítems.

Hasta aquí nuestra exposición ha abarcado la descripción de la técnica de elaboración de una escala tipo Guttman y de la evaluación de su grado de adecuación al modelo teórico. Las presentaciones de la técnica de Guttman generalmente se detienen en este punto. Vale la pena, sin embargo, dar un paso más e incursionar en un aspecto de aplicación menos frecuente de este tipo de escala.

4

Medición del grado de "disposición al cambio" de la muestra de empresarios: dirección e intensidad

La aplicación de la escala al grupo de empresarios entrevistados indicó la distribución contenida en el cuadro 6. Los resultados

¹⁴ Lamentablemente Guttman no ha proporcionado un criterio objetivo para determinar con precisión cuál es la frecuencia de pautas fuera de escala que indica la presencia de errores de naturaleza sistemática.

son claros: la mayoría de los empresarios se situó en los puntos intermedios de la escala. Solo muy escasas minorías dieron evidencias de una extrema aceptación o de un extremo rechazo hacia el cambio. Como puede verse en la columna de porcentaje acumulado, la mitad de los entrevistados recibió un puntaje de "3".

Cuadro 6

Distribución de los puntajes asignados a la muestra de empresarios en la escala de "disposición al cambio"

	Puntaje	Frecuencia empresarios	Porcentaje empresarios	Porcentaje acumulado
baja disposición	0	3	3	3
	1	5	5	8
	2	19	18	26
	3	28	26	52
	4	46	43	95
alta disposición	5	6	5	100
		(107)	100 %	

Los resultados obtenidos provienen del ordenamiento de los entrevistados en términos de la frecuencia de respuestas positivas dadas a la batería de ítems. Este ordenamiento permite decir cuál es la proporción de entrevistados que obtuvo un puntaje de "5", un puntaje de "4", etcétera. Pero este ordenamiento —en términos de la frecuencia de respuestas positivas— no informa acerca de cuál es la posición en la escala que implica, para la muestra estudiada, un cambio de dirección de la variable latente, es decir, un cambio de respuesta favorable a desfavorable hacia el cambio. Para decirlo de otro modo, un puntaje de "5" indica un mayor grado de disposición hacia el cambio que un puntaje de "4", y un puntaje de "1", un mayor grado que uno de "0". Pero, ¿qué sector de puntajes indica, independientemente del grado (intensidad), una disposición positiva y cuál una disposición negativa (dirección)?

Este problema no es privativo de las escalas acumulativas. Tampoco las categorías de una escala tipo Likert informan inequívocamente acerca de la dirección de la actitud que expresan.

Es el investigador quien debe decidir en cada caso dónde se ubica el punto de corte direccional entre las categorías "muy de acuerdo", "de acuerdo", "indiferente", "en desacuerdo" y "muy en desacuerdo" (o sus equivalentes).

Guttman y Suchman (1947b) han sugerido un procedimiento para determinar "de manera objetiva e inequívoca" cuál es la zona favorable, y cuál la desfavorable de una escala que mida una variable de actitud y que haya probado adecuarse al modelo de unidimensionalidad. El procedimiento se basa sobre el supuesto de que los individuos que responden frecuentemente "sí" o "no", es decir, que reciben puntajes extremos en la escala, tienden a tener sentimientos extremos, a sentir intensamente acerca del área psicológica en cuestión. Son quienes —dependiendo del particular fraseo usado por el investigador— responden frecuentemente "decididamente", "absolutamente", "indudablemente" (sí o no). Por otra parte, se supone que quienes no dan una alta frecuencia de respuesta positiva o negativa y que, por lo tanto, reciben puntajes intermedios en la escala, tienden a ser más indiferentes o a tener sentimientos menos intensos respecto del área psicológica en cuestión. Son quienes —también con las correspondientes variaciones de fraseo— responden "indeciso", "probablemente", "preferentemente" (sí o no). En otras palabras, el supuesto básico es que la intensidad de sentimientos es mayor en ambos extremos de la escala y que decrece hacia el medio.

La actitud de un individuo hacia un área psicológica cualquiera puede medirse, entonces, en términos de su contenido (dirección) y de su intensidad, es decir, en términos de sus contenidos cognitivos y afectivos. La correlación entre los puntajes de contenido y los de intensidad obtenidos para una serie de muestras y una serie de variables de actitud ha probado dar curvas en forma de U y de J. Es decir, a medida que aumenta la distancia del polo positivo de la escala de contenido hacia el centro, la intensidad de los sentimientos decrece y llega a un punto (o zona) en que empieza a crecer hasta alcanzar nuevamente un máximo en el extremo negativo de la escala de contenido. El punto (o zona) más baja de la curva representa la posición en la escala de contenido que ocupan los individuos con sentimientos menos intensos. Este punto (o zona) es considerado el de indiferencia o neutralidad afectiva.

Si la correlación fuera perfecta, habría una relación curvilínea perfecta, es decir, todos los individuos con un mismo puntaje de contenido tendrían el mismo puntaje de intensidad. Si no hubiera errores, sería entonces posible determinar con exactitud el punto de cambio de dirección del continuum de actitud —allí donde se ubica el puntaje de contenido caracterizado por haber recibido el menor puntaje de intensidad de la muestra. Dicho punto (o zona) se supone invariante, es decir, independiente de los ítems elegidos y de su particular fraseo. Dado que habitualmente hay error —que individuos con el mismo puntaje de contenido pueden tener diferentes puntajes de intensidad—, Guttman y Suchman han sugerido trabajar con la mediana de intensidad de cada puntaje de contenido y usar la curva de tales medianas como aproximación a la verdadera función de intensidad.

Para ordenar a los individuos a lo largo de la escala de intensidad de una manera independiente de su ordenamiento a lo largo de la escala de contenido, se ha sugerido preguntar a continuación de haber obtenido las respuestas de contenido, “¿cuán seguro (convencido, creído, etcétera) está usted de lo dicho?”. Las opciones (categorías de respuesta) a esta pregunta reciben un puntaje, de modo que cada individuo obtiene finalmente dos puntajes por cada ítem respondido, uno de contenido y uno de intensidad. La tabulación cruzada de ambos puntajes provee los datos para construir la curva de intensidad. Como se ha dicho, esta es la curva de las medianas de intensidad de cada puntaje de contenido. Para determinar dichas medianas es necesario identificar el percentil de intensidad que le corresponde en la muestra al individuo que, para cada puntaje de contenido, ocupa la posición media de intensidad, cálculo que se hace por interpolación.

A continuación —en el cuadro 7 y gráfico 1— presentaremos el resultado de la aplicación del procedimiento descrito a los datos producidos por la escala de “disposición al cambio”. (Vale la pena insistir en que su aplicación fue posible una vez que se determinó que los datos indicaban la presencia de una escala unidimensional. De no haber sido así su aplicación hubiera carecido de sentido, puesto que no se habrían satisfecho los supuestos del procedimiento.)

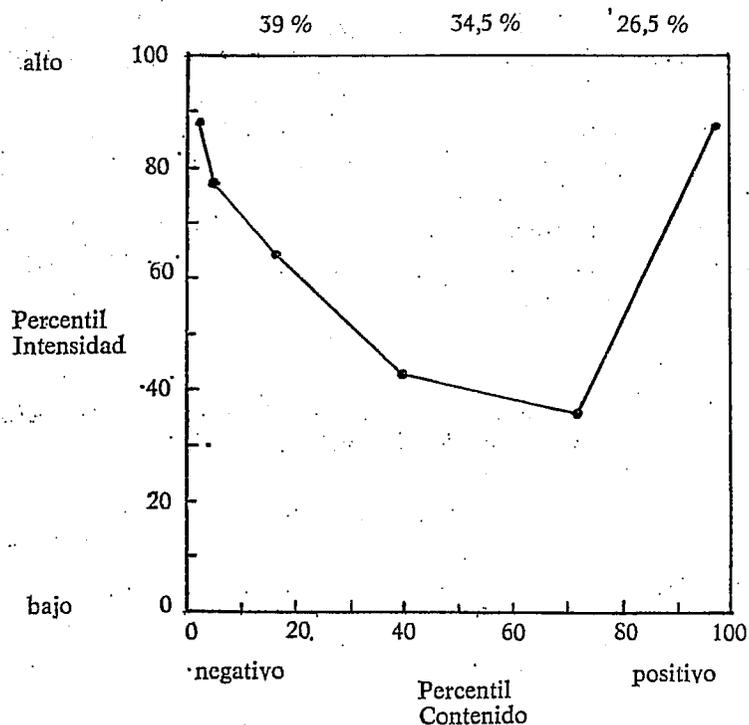
Como se recordará —cf. nota al pie 3 y p. 382— luego de obtener las respuestas de los entrevistados a cada uno de los nueve ítems, se les preguntó por las razones a las que había obedecido su elección. El análisis de este material —registrado *verbatim*— sirvió para verificar el contenido e intensidad de las respuestas dadas a los nueve ítems y codificadas como “1”, “2”, “3” y “4”. Se asignó un puntaje de “0” a cada respuesta de poca intensidad afectiva (codificadas como “2” y “3”) y uno de “1” a las de mucha intensidad (codificadas como “1” y “4”). La amplitud del puntaje de intensidad varía, en consecuencia, entre “0” y “5”, igual que el puntaje de contenido.

Cuadro 7

Correlación entre los puntajes de contenido y los puntajes de intensidad en la escala de “disposición al cambio”

Puntaje de intensidad	n	Puntaje de contenido						n total	%
		(neg.)			(pos.)				
		0	1	2	3	4	5		
(alto) 5	5	2	2	4	5	2	4	19	100
4	4	1	1	3	2	7	—	14	82
3	3	—	1	7	3	3	1	15	69
2	2	—	1	4	9	15	1	30	55
1	1	—	—	1	5	11	—	17	27
(bajo) 0	0	—	—	—	4	8	—	12	11
n total		3	5	19	28	46	6	107	
% acum.		3	8	26	52	95	100		
punto medio percentiles contenido		1,5	5,5	17	39	73,5	97,5		
mediana percentiles intensidad		86,5	75,5	64	42,5	34,5	86,5		

Gráfico 1



Representación gráfica de la correlación entre los puntajes de contenido y los de intensidad

El cuadro 7 resume los resultados de la tabulación cruzada de los puntajes de contenido por los de intensidad. Contiene, además, los datos necesarios para obtener la representación de la curva de medianas de intensidad que ilustra el gráfico 1.

El cálculo del punto medio de los percentiles de contenido es simple. Se obtiene sobre la base de la fila de porcentajes acumulados de los puntajes de contenido. El correspondiente al puntaje "0" es igual a 1,5, es decir, a la mitad de su porcentaje acumulado. El correspondiente a cada uno de los puntajes restantes es igual a la suma del porcentaje acumulado del puntaje inmediato inferior más la mitad de la diferencia entre el porcentaje acumulado del puntaje que interesa menos el del inmediato

inferior. Por ejemplo, el punto medio del percentil de contenido correspondiente al puntaje "3" es igual a

$$26 + \frac{52 - 26}{2} = 39$$

El cálculo de la mediana de los percentiles de intensidad se obtiene mediante la aplicación de la fórmula de interpolación de la mediana para una distribución continua en la que los datos están agrupados en intervalos de clase. Consideremos, por ejemplo, el cálculo de la mediana del percentil de intensidad correspondiente al puntaje de contenido "4". Habrá de notarse que dicha mediana —expresada en percentiles— está contenida en el intervalo 27 a 55, correspondiente a los puntajes de intensidad "1" a "2". Por lo tanto, la mediana buscada debe ser superior a 27 e inferior a 55. La fórmula general de la mediana es igual a:

$$Md = l_{Md} + \left[\left(\frac{N/2 - f_{Md \text{ acum.}}}{f_{Md}} \right) (i) \right]$$

En el caso del puntaje de contenido "4",

- l_{Md} = límite inferior del intervalo que contiene la mediana.
- $N/2$ = mitad de la frecuencia total de sujetos con puntaje "4".
- $f_{Md \text{ acum.}}$ = suma de las frecuencias de los sujetos con puntaje "4" hasta el límite inferior del intervalo que contiene la mediana de dicho puntaje (8 + 11).
- f_{Md} = frecuencia de sujetos en el intervalo que contiene la mediana del puntaje "4".
- i = tamaño del intervalo de clase que contiene la mediana.

Reemplazando en la fórmula general, la mediana del percentil de intensidad correspondiente al puntaje de contenido "4" es igual a:

$$Md^{4} = 27 + \left[\left(\frac{46/2 - 19}{15} \right) (55 - 27) \right] = 34,46 \cong 34,5$$

La curva en forma de "U" obtenida al volcar sobre el gráfico las medianas de intensidad de los puntajes de contenido reveló —como puede observarse en el gráfico 1— una zona relativamente horizontal en su parte inferior. Esto habitualmente dificulta la determinación de la posición exacta de un punto neutro definido. Permite, sin embargo, detectar tres subgrupos, dos de ellos con actitudes muy definidas —en pro o en contra— y un tercero relativamente indiferente. En nuestro caso, los resultados hicieron posible concluir que el porcentaje de la muestra investigada que reveló poseer una disposición decididamente favorable al cambio en el desempeño del rol de empresario es inferior al del que reveló una disposición decididamente desfavorable. El primer subgrupo concentra —por lo menos— 26,5 % de la muestra y el segundo, 39 %. El subgrupo calificado de relativamente indiferente reúne alrededor de un tercio de la muestra, 34,5 % como máximo.

5

Consideraciones finales

En este capítulo se expuso en detalle el proceso seguido en la elaboración de una escala de actitud tipo Guttman. Comenzando por la delimitación de la dimensión psicológica que interesaba medir, hemos recorrido el camino que pasa por la elaboración de los ítems, la recolección y procesamiento de las respuestas a los mismos y la evaluación del grado de adecuación del instrumento al modelo teórico, para desembocar en la aplicación de dicho instrumento a una muestra de individuos. Antes de concluir este trabajo, se hace necesario ubicar su contenido dentro del contexto más general de la utilización de escalas en la tarea de investigación.

Lo que aquí se ha expuesto está contenido estrictamente dentro de los límites de la construcción de un instrumento de medición y de la evaluación de su adecuación al modelo teórico. Pero la evaluación de tal adecuación no constituye el fin del camino. Antes de proceder a la adopción del instrumento resta

evaluar su grado de validez y de confiabilidad. En efecto, la obtención de un coeficiente de reproducibilidad adecuado, de marginales variados y no extremos, de una distribución aleatoria de las pautas de respuesta fuera de escala, son todos criterios que miden la adecuación del instrumento al modelo de Guttman. Otro problema muy distinto es evaluar el grado en que el instrumento diseñado mide efectivamente la dimensión pretendida, y no alguna otra, y el grado en que aplicaciones repetidas de manera independiente producen resultados consistentes. Esta etapa —en la que no hemos incursionado en este trabajo y cuyo proceso ha sido tratado con extraordinaria lucidez en los capítulos de Cronbach y Meehl y de Thorndike incluidos en este volumen— debe ser recorrida antes de aceptar el instrumento e intentar aplicarlo en la investigación.

Referencias bibliográficas

- Edwards, A. L. (1957), *Techniques of Attitude Scale Construction*, Appleton-Century-Crofts, Nueva York.
- Ford, R. N. (1950), "A Rapid Scoring Procedure for Scaling Attitude Questions", *Public Opinion Quarterly*, nº 14, pp. 507-532.
- Goodenough, W. H. (1944), "A Technique for Scale Analysis", *Educational Psychological Measurement*, nº 4, pp. 179-190.
- Guttman, L. (1947), "The Cornell Technique for Scale and Intensity Analysis", *Educational Psychological Measurement*, nº 7, pp. 247-280.
- Guttman, L. y E. A. Suchman (1947b), "Intensity and a Zero Point for Attitude Analysis", *American Sociological Review*, nº 12, pp. 55-67.
- Kogan, N. y M. A. Wallach (1964), *Risk Taking: A Study in Cognition and Personality*, Holt, Rinehart & Winston, Nueva York.
- Noland, E. W. (1945), "Worker Attitude and Industrial Absenteeism: A Statistical Appraisal", *American Sociological Review*, nº 10, pp. 503-510.
- Sautu, R. y C. Wainerman (1971), *El empresario y la innovación*, Editorial del Instituto, Buenos Aires.
- Suchman, E. A. (1950), "The Scalogram Board Technique for Scale Analysis", en S. A. Stouffer *et al.* (compil.), *Measurement and Prediction*, Princeton University Press, cap. 4, Princeton, N. J.
- Toby, J. y M. L. Toby (1954), "A Method of Selecting Dichotomous Items by Cross Tabulation", en M. W. Riley *et al.* (compil.), *Sociological Studies in Scale Analysis*, Rutgers University Press, N. J.