5

Procedimientos de captación de Información: Tareas de Elección y Tareas de Ordenación.

Un método de obtención de información es un procedimiento o tarea por el cual obtenemos el juicio o respuesta de un sujeto ante un ítem o un conjunto de items. El método de obtención de información acaba en la respuesta o juicio del sujeto ante la tarea.

Cómo se valore esa respuesta o juicio y qué se haga con las puntuaciones hasta obtener la escala es ya cuestión del procedimiento de escalamiento que se utilice. Ese es precisamente el cometido de un procedimiento de escalamiento: convertir la respuesta o juicio de un sujeto ante una tarea en valores de escala.

Los datos recogidos por un mismo método o tarea de obtención de información pueden utilizarse con distintos procedimientos de escalamiento. Lo inverso también es cierto: un mismo procedimiento de escalamiento puede utilizar diferentes métodos de obtención de información como punto de partida. Sin embargo, no todos los procedimientos de escalamiento pueden utilizarse con todos los métodos de obtención de información.

Debe quedar claro, pues, que los métodos de obtención de información y los procedimientos de escalamiento son piezas distintas dentro de un método de escalamiento en sentido global.

En esta unidad nos vamos a ocupar de los métodos o tareas de obtención de información. Ahora bien, para facilitar su conexión con las unidades siguientes relativas a métodos de escalamiento hemos introducido en la discusión de cada tarea de obtención de información un apartado sobre cómo aprovechar la información ofrecida por esa tarea para escalar los items.

Supongamos que tenemos 10 enunciados (A, B, C, D, E, F, G, H, I, y J) referidos al tema del aborto. Deseamos obtener información de una muestra de sujetos acerca de cuán favorables al aborto (y por tanto cuán desfavorables) les parecen (Orientación de Juicio). O bien, deseamos obtener información de una muestra de sujetos acerca de cuán favorables o desfavorables son los sujetos al aborto (Orientación de Respuesta). En ambos casos se puede utilizar una serie de diferentes tareas para captar la información necesaria. Todas las tareas de captación de información que se presentan a continuación pueden utilizarse para obtener tanto "juicios" como "respuestas", si la naturaleza del tema lo permite. En cada caso expondremos algún ejemplo de la orientación de juicio; el lector puede fácilmente convertir cada caso a la orientación de respuesta.

La distinción entre el grupo de tareas o métodos de elección y el grupo de tareas o métodos de ordenación se ofrecerá al final, como síntesis ordenadora de los seis métodos concretos de captación de información que se exponen a continuación. Sin embargo, puede anticiparse que los cuatro primeros métodos de los que nos ocuparemos son métodos de elección, mientras que los dos últimos son métodos de ordenación.

5.1

1. Tareas de estímulo simple.

En una tarea de estímulo simple se presenta al sujeto un solo ítem y se le solicita que responda si cumple o no una condición.

Siguiendo con nuestro ejemplo, podemos presentarle al sujeto el enunciado B y preguntarle a continuación:

"¿Considera Ud. que esta opinión es favorable al aborto?"

a) "Sí" b) "No"

¿Cómo aprovechar este método de obtención de información para escalar los items? De esta tarea de estímulo simple (juicio) puede derivarse fácilmente información que permita escalar los items. Si tenemos una muestra de sujetos, cada uno de los cuales realiza esta tarea para cada ítem, entonces para cada ítem tendremos la frecuencia de sujetos que ha dicho "Sí". Resulta claro que pueden ordenarse los items de más (más frecuencia de "síes") a menos favorable al aborto, y utilizar la frecuencia de "síes" como información para determinar las puntuaciones escalares de los items.

5.2

2. *Tareas de elección de alternativas.*

En una tarea de elección de alternativas se presenta al sujeto un solo ítem y se le solicita que elija aquella alternativa, de entre un conjunto prefijado, que mejor describa el grado en que el ítem cumple una condición.

En nuestro ejemplo podemos presentarle al sujeto el estímulo B y preguntarle:

> "¿En qué medida considera Ud. que este enunciado representa una opinión favorable al aborto?" Escoja una de las alternativas de respuesta siguientes.
>
> 1) Nada favorable
>
> 2) Poco favorable
>
> 3) Bastante favorable
>
> 4) Muy favorable.

El sujeto está calificando cada ítem con una puntuación de 1 a 4 en función de la "favorabilidad" al aborto de cada ítem. En este ejemplo concreto, cada sujeto da a cada ítem una puntuación de 1 a 4. El texto que acompaña o explica cada grado se denomina a veces "anclaje" o "anclaje verbal", para indicar un lugar de la escala al que se ha pretendido dar significado en lenguaje común.

¿Cómo aprovechar este método de obtención de información para escalar los items? No es difícil intuir cómo se puede aprovechar esta información para escalar los 10 items. Si sumamos los puntos que ha obtenido cada ítem a través de todos los sujetos podremos clasificar los items del más favorable (puntuación más alta) al menos favorable, y estas puntuaciones ofrecerán la información de partida para determinar los valores de escala. Una variante de este método consiste en promediar la puntuación que ha obtenido cada ítem a través de todos los sujetos, y utilizar estos promedios para obtener el valor de escala.

Variante de asignación libre de puntuación dentro de un rango. Hay una variante interesante de este método que a veces se utiliza tanto con fines de juicio como de respuesta. Consiste en permitir que el sujeto asigne al estímulo no sólo los valores enteros que acompañan a las alternativas, sino cualquier valor decimal comprendido entre el máximo y el mínimo de la escala. De este modo se pretende que el sujeto pueda expresar de un modo más preciso su evaluación del estímulo.

A veces el procedimiento se vuelve más abierto suprimiendo las "alternativas" enteras y sus textos. Simplemente se pide al sujeto que asigne un número entre un máximo y un mínimo.

Si efectivamente el sujeto es capaz de efectuar discriminaciones tan finas, el método permite la expresión de las mismas. Sin embargo, el uso de un máximo y mínimo sin textos que actúen como anclajes vuelve todavía más ambigua y subjetiva la interpretación de la escala de respuesta.

5.3

3. Comparación de estímulos por pares.

Podemos formar todos los pares posibles entre un conjunto limitado de estímulos, presentarle a cada sujeto cada par de estímulos, y preguntarle cuál de ellos tiene más (o menos) de una cualidad.

Por ejemplo, supongamos que tenemos los tres estímulos A, B y C, podemos formar todos los pares posibles entre ellos:

AB BC

AC

Le presentamos el primer par de enunciados (el par A-B) a un sujeto y le preguntamos:

> "¿Cuál de estos dos estímulos es más favorable al aborto?"

Después le presentamos el siguiente par y le volvemos a hacer la misma pregunta. Y así con todos los pares. Después se repite el mismo proceso con cada sujeto.

¿Cómo aprovechar la información de este método para escalar los items? Cada estímulo aparece en el mismo número de pares. Si hay n estímulos, cada estímulo aparece en n-1 pares. Es decir, un estímulo concreto aparece en un par con cada estímulo que no es él mismo. Por tanto, si un sujeto dice que un estímulo concreto es el "mayor" en cada par en que aparece ese estímulo, como máximo lo habrá escogido como el mayor n-1 veces, cada vez que se le ha presentado en un par.

Si un estímulo A ha sido escogido por un sujeto n - 1 veces, ningún otro estímulo B puede ser escogido n - 1 veces. Dado que el sujeto sólo puede escoger, y ha de escoger, necesariamente, un estímulo de cada par; y dado que hay uno y sólo un par (A, B), entonces, un estímulo B puede ser escogido en el caso máximo en n-2 casos. Si se generaliza este razonamiento se encontrará que el método permite obtener, para cada sujeto, un conjunto de elecciones graduadas. (No obstante no siempre los sujetos deciden en realidad con un esquema perfectamente lógico, lo cual puede ser objeto de contraste).

Si hay N sujetos, tenemos N(n-1) informaciones sobre cada estímulo. Para escalar los items a partir de esta información bastará con contar el número de veces que cada estímulo ha sido preferido a los demás. Por ejemplo, el estímulo A: ¿Cuántas veces ha sido elegido como más favorable al aborto? Como máximo el estímulo A puede haber sido elegido N(n-1) veces como el más favorable al aborto, (en el supuesto de que fuera el estímulo más favorable al aborto y todos los sujetos estuvieran de acuerdo en ello), y como mínimo 0 veces, (en el supuesto de que fuera el menos favorable al aborto y los N sujetos que actúan como jueces estuvieran de acuerdo en ello). De este modo, después de una tarea de comparación de estímulos por pares, cada estímulo ha recibido una puntuación entre 0 y N(n-1). Esa puntuación permite fácilmente escalar los items.

Número de pares. El método de obtención de información por comparación por pares de estímulos presenta la dificultad práctica de que el número de pares a presentar a cada sujeto crece muy de prisa al incrementarse el número de estímulos a escalar.

En general, hacen falta **p** pares para **n** estímulos, siendo p:

$$p = \binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2} = \frac{n^2-n}{2}$$

Para 3 estímulos sólo hacen falta 3 pares, pero para 10 ya son necesarios 45 pares. Por ello rara vez resulta práctico utilizar este método de captación de información con más de 10 estímulos.

Cuando el número de estímulos es grande también es posible obtener todos los pares pero no presentar todos los pares a los sujetos. Se puede presentar una muestra, aleatoria o compensada, de pares. Puede utilizarse la misma muestra para todos los sujetos o diferentes muestras aleatorias para todos o parte de los sujetos. Estas posibilidades, cuyo uso no es muy frecuente, permiten también obtener para cada estímulo el número total de veces que ha sido preferido a otro, y, a partir de ese total operar para obtener el valor de escala.

Controles de consistencia de los sujetos. Estrictamente hablando, un par es conjunto de dos elementos donde importa el orden. En ese sentido con el conjunto C formado por los estímulos A y B pueden formarse los pares {A, A}, {A,B}, {B,A} y {B,B}, que son todos los elementos del producto cartesiano CxC. Sin embargo, como se habrá observado, aquí hemos utilizado el término par de un modo diferente, para designar cualquier combinación sin repetición de los estímulos. En este uso el par {A,A} no tiene sentido, dado que no tiene sentido preguntarle a un sujeto si el estímulo A tiene más o menos de una cualidad que el estímulo A. En este uso, además, el par {A,B} se considera el mismo que el par {B,A}. Si los sujetos responden de un modo lógico y sin errores, entonces se puede suponer que si escogen A como mayor que B en el par {A,B} harían lo mismo en el par {B,A}. En la práctica, esto puede no ser siempre así. Algunas veces puede resultar de interés poner a prueba esta hipótesis ofreciendo ambos pares -en posiciones distanciadas- para que el sujeto los juzgue. Si en una tarea de escalamiento se introducen los dos pares que contienen los mismos dos elementos, se puede contrastar -salvo efectos de memoria- la consistencia de los jueces.

Por otra parte, si tenemos tres estímulos A, B y C y un sujeto ha afirmado que A›B, y que B›C, entonces debería afirmar que A›C. Sin embargo, es posible que un sujeto responda C›A; en este caso no queda claro cuál de los tres estímulos es el que presenta en mayor grado la cualidad. Cuando esto último sucede se dice que se ha formado una *tríada circular*. El número de tríadas circulares que aparecen en unos datos sirve para poner de manifiesto el grado de consistencia de los sujetos.

Algunas veces si unos estímulos presentan inconsistencias, mediante el control de pares o mediante el control de tríadas, ello puede ser atribuído a que los estímulos no son escalables, en lugar de atribuirlo a que los sujetos no son consistentes en sus juicios o en sus respuestas. En realidad, lo único que tenemos si ello sucede son unos datos inconsistentes que pueden deberse a que los sujetos no son consistentes, a que los estímulos no son escalables, a otros defectos prácticos -por ejemplo, los sujetos no entendieron bien las instrucciones o estuvieron distraídos-, o a una combinación de estas razones. La repetición sistemática del proceso (variando los sujetos, variando los estímulos) puede ofrecer datos acerca de cual es la razón de la dificultad.

Comparación de estímulos por pares y método de las comparaciones apareadas de Thurstone. El método de obtención de información "comparación de estímulos por pares" a veces se denomina abreviadamente 'método de comparaciones apareadas'. Este uso puede inducir a cierta confusión. El "método de comparaciones apareadas" es el nombre de un método de escalamiento *en su totalidad,* creado por Thurstone. El método de comparaciones apareadas de Thurstone utiliza, efectivamente, la comparación de estímulos por pares como método de captación de información, pero como un método de escalamiento que es, asume además determinados supuestos y procesos particulares para escalar los items. Por eso, no puede confundirse el "método de las comparaciones apareadas de Thurstone", que es un método de escalamiento con todos sus pasos hasta la medición de sujetos, con el método o tarea de "comparación de estímulos por pares", que es tan solo un método de obtención de información.

El método de comparación de estímulos por pares puede utilizarse con orientación de juicio o con orientación de respuesta, y puede servir como fuente de obtención de información a diversos métodos de escalamiento, para que estos puedan obtener de diferentes formas las puntuaciones de escala.

5.4

4. *Comparación de estímulos por grupos.*

En este procedimiento, análogo al de comparación de estímulos por pares, se ordenan los estímulos en grupos de tres, o más, estímulos y se solicita al sujeto que escoja en cada grupo aquel estímulo que presenta más (o menos) de una cualidad.

Por ejemplo, podemos agrupar 5 enunciados, con contenidos sobre el tema de la discriminación racial, en todas las combinaciones posibles de 3 enunciados, y pedirle al sujeto que elija cual de los tres enunciados de cada grupo enuncia una idea más xenófoba. En este caso se trataría de una tarea de comparación de estímulos por tríos, orientada a juicio.

Sean los enunciados:

> A. Todas las personas han de ser iguales ante la ley, sin distinciones por razones de sexo, raza o religión.
>
> B. Debería limitarse el número de emigrantes extranjeros que acepta la Comunidad Europea, de lo contrario no habrá trabajo para los europeos.
>
> C. Un extranjero sin dinero y sin trabajo debe ser inmediatamente expulsado.
>
> D. Los servicios sociales deberían atender preferentemente a los extranjeros sin recursos, sin trabajo o sin vivienda.

E. Deberían arbitrarse fórmulas legales para limitar la estancia de extranjeros emigrantes, protegiendo sus derechos y los de los trabajadores locales.

Se forman todos las combinaciones posibles de tres enunciados:

ABC	BCD	CDE
ABD	BCE	
ABE	BDE	
ACD		
ACE		
ADE		

En total, diez combinaciones o grupos de tres enunciados. Ahora se le presentan a cada sujeto que actúa como juez cada una de estas diez combinaciones, y, ante cada trío de enunciados se le pregunta:

¿Cuál de estos tres enunciados cree Ud. que expresa una idea más contraria a los extranjeros?

El sujeto produce una elección por cada trío. Por ejemplo, ante el primer trío de enunciados, ABC, un sujeto puede responder que el enunciado C es el más xenófobo. Esto puede representarse abreviadamente así:

C›(A ó B)

El sujeto emite tantos juicios (o da tantas respuestas) como tríos se le presentan; diez en este caso concreto.

Obsérvese que el procedimiento puede también realizarse con grupos de 4 estímulos, de 5 estímulos, de 6, etc. El tamaño del grupo de estímulos que se presentan juntos al sujeto, que denominaremos k, puede variar entre 3 y n, siendo n el número total de estímulos a escalar:

$$3 \leq k \leq n$$

Número de grupos en que aparece un estímulo. Cada estímulo aparece en el mismo número de grupos. En el ejemplo anterior cada enunciado aparecía en 6 tríos. Dados n estímulos, que queremos presentar en todos los grupos posibles de k estímulos, cada estímulo estará en tantos grupos como combinaciones de k-1 estímulos puedan formarse con n - 1 estímulos.

Es decir, en general, el número de combinaciones en que aparece un estímulo es:

$$\binom{n-1}{k-1} = \frac{(n-1)!}{(k-1)!\,[(n-1)-(k-1)]!} = \frac{(n-1)!}{(k-1)!\,(n-k)!}$$

Para k = 2, caso que estudiamos en la comparación de estímulos por pares, tendremos, como ya vimos:

$$\binom{n-1}{2-1} = \binom{n-1}{1} = \frac{(n-1)!}{1!\,(n-1-1)!} = \frac{(n-1)!}{(n-2)!} = n-1$$

Por ejemplo, en el caso que acabamos de ver con cinco enunciados (n=5) tomados en combinaciones de 3 (k=3) tendríamos:

$$\binom{n-1}{k-1} = \binom{5-1}{3-1} = \binom{4}{2} = \frac{4!}{2!\,(4-2)!} = 6$$

como hemos podido comprobar.

Puntuación mínima y máxima que puede recibir un estímulo.

Si cada estímulo está en $\binom{n-1}{k-1}$ grupos, entonces, después de haberle presentado todos los grupos a un sujeto, éste habrá podido escoger como máximo $\binom{n-1}{k-1}$ veces a ese estímulo como el "mayor" en la cualidad solicitada.

Considerando cada elección como un punto, un estímulo determinado puede recibir de 0 a $\binom{n-1}{k-1}$ puntos por cada sujeto.

Si los k grupos de estímulos son presentados completos a N sujetos, entonces la puntuación que un estímulo puede recibir en total varía entre 0 y $N\binom{n-1}{k-1}$.

¿Cómo aprovechar los datos ofrecidos por este método de obtención de información para escalar los items? Para escalar los items a partir de esta información bastará con contar el número de veces que cada estímulo ha sido preferido a los demás.

Por ejemplo, el estímulo C: ¿Cuántas veces ha sido elegido como más xenófobo en cada tríada, (o tétrada, etc.)? Como máximo el estímulo C puede haber sido elegido $N\binom{n-1}{k-1}$ veces como el más xenófobo, (en el supuesto de que fuera el estímulo más xenófobo y de que todos los sujetos estuvieran de acuerdo en ello), y como mínimo 0 veces, (en el supuesto de que fuera el menos xenófobo y los N jueces estuvieran de acuerdo en ello).

Después de una tarea de comparación de estímulos por grupos, cada estímulo ha recibido una puntuación entre 0 y $N\binom{n-1}{k-1}$. Esas puntuaciones dan idea de en qué orden los estímulos son considerados como más o menos xenófobos y también de qué "distancia" hay entre los estímulos, según la percepción de nuestra muestra de sujetos. Esas puntuaciones son las que permiten fácilmente escalar los estímulos.

Número de grupos. El número de grupos a formar depende del número de estímulos a escalar y del tamaño de los grupos. En general, cuánto menor sea el número de estímulos a escalar n y más próximo al mismo esté el tamaño de los grupos k, menor número de grupos hay que formar.

En general, hacen falta **g** grupos o combinaciones para n estímulos, siendo g:

$$g = \binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

Variantes del método respecto al número de grupos presentados a cada sujeto. Cuando el número de estímulos es grande, es posible obtener todos los grupos pero no presentar todos los grupos a los sujetos. Se puede presentar una muestra, aleatoria o compensada, de esos g grupos. Puede utilizarse la misma muestra de grupos para todos los sujetos o diferentes muestras aleatorias, sucesivas u ordenadas, para todos o parte de los sujetos.

Por ejemplo, si tenemos un valor g muy alto, es posible presentar a los sujetos una parte del total de grupos en un proceso cíclico. Es decir, los 10 primeros grupos se presentan al sujeto 1, los 10 segundos al 2, etc. Cuando se terminan los pares se vuelve a empezar. Fijado el N de la muestra, el número de grupos que se presenta a cada sujeto, número que denominaremos **d**, determina el número de sujetos que juzga cada estímulo. El valor d se escoge de modo que sea un valor razonable y que g/d sea exacto (si es posible). El valor N/(g/d) también debe ser exacto para que cada grupo, y por tanto cada estímulo, haya sido enjuiciado, al final, el mismo número de veces. Si se utiliza esta posibilidad, entonces la puntuación de cada estímulo puede estar entre 0 y $Nd\binom{n-1}{k-1}/g$.

Estas posibilidades, cuyo uso no es muy frecuente, permiten también obtener para cada estímulo el número total de veces que ha sido preferido a otros, y, a partir de ese total, operar para obtener el valor de escala.

Además, por supuesto, los controles de consistencia explicados para la comparación de estímulos por pares, pueden extenderse adaptados a la comparación de estímulos por grupos.

El caso particular $k = n$. Si k, el número de estímulos que se presentan en cada grupo, es igual a n, el número total de estímulos, entonces, obviamente, la tarea se reduce a presentar todos los estímulos a la vez y preguntarle al sujeto cuál de ellos presenta más (o menos) de una cualidad bajo consideración. Este caso tiene muy poco interés práctico en cuanto que no es un método muy utilizado; pero, como veremos enseguida, tiene mucho interés teórico por cuanto nos ayuda a entender simplificadamente algunas de las cuestiones más interesantes y sorprendentes de los métodos de escalamiento.

Si tenemos N sujetos a los que se presenta una tarea de comparación de estímulos con k=n, entonces la puntuación de un estímulo concreto puede variar entre 0 (ningún sujeto lo escoge como el mayor en la dimensión) y N (todos los sujetos lo escogen como el mayor en la dimensión). Una tarea de comparación con k=n simplifica el trabajo de los sujetos y el cómputo de puntos recibidos por cada estímulo; pero, a su vez, reduce la información disponible para escalar los estímulos.

Obsérvese que si los N sujetos coinciden en considerar a un estímulo determinado como el que posee en más grado la cualidad bajo consideración, la puntuación de ese estímulo en este método será N, pero todos los demás estímulos bajo consideración sin distinción presentarán una puntuación 0, porque no habrán sido escogidos ni una sola vez. Ello pone de manifiesto con claridad que cuando k tiende a n el proceso tiende a volverse más sencillo (menos grupos que presentar a cada sujeto) pero el resultado tiende a ser más tosco, (menos oportunidades tienen los sujetos de mostrar los matices de sus preferencias). Cuando k tiende a n el método tiende a perder precisión, a ofrecer menos información.

Número de grados de desigualdad que pueden extraerse de n estímulos presentados a elección en grupos de k. Si tenemos cuatro elementos A,B,C y D y los presentamos en un solo grupo (k=n=4), un sujeto sólo puede establecer o manifestar la desigualdad entre un estímulo y todos los demás. Por ejemplo, puede decir que el estímulo A es el que presenta más cualidad, el más alto en la dimensión:

A›(B,C ó D)

Aquí sólo se ha obtenido un grado de desigualdad, una relación de desigualdad, sólo se ha establecido un escalón o peldaño. Sabemos que A es considerado "mayor" que cualquiera de los otros tres estímulos, pero no sabemos que relación hay entre B y C, entre B y D o entre C y D; sólo sabemos que cualquiera de los tres está por debajo de A en la dimensión bajo consideración. Así podemos decir que con 4 estímulos y k=4 sólo hemos podido establecer un grado de desigualdad.

Si 4 estímulos los presentamos en grupos de 3 estímulos (k=3), un sujeto consistente sólo puede establecer dos grados de desigualdad, por ejemplo:

A›B›(C ó D)

Si los presentamos por pares (k=2), un sujeto consistente puede manifestar la desigualdad entre todos los estímulos, por ejemplo:

A›B›C›D

Estableciendo tres grados de desigualdad, que es el máximo de desigualdades encadenadas que pueden realizarse con cuatro elementos.

En general, con n elementos se pueden establecer como máximo n-1 grados de desigualdad.

Con n=5 y k=5 se puede establecer un grado de desigualdad; con n =5 y k=4 se pueden establecer dos grados de desigualdad; con k=3 se pueden establecer tres; con k=2, cuatro, es decir, todos los posibles entre cinco estímulos.

En general, con n estímulos a escalar y k estímulos en cada grupo, siendo $2 \leq k \leq n$, se pueden establecer consistentemente n-k+1 grados de desigualdad.

Esos grados de desigualdad son la información ordinal que aporta el método de captación de información, por eso la tarea con k=n es la que da un resultado más pobre y la tarea con k=2 aquella que permite obtener el orden completo de los estímulos.

El hecho de que cuando k › 2 no se pueda obtener el orden completo de estímulos hace este método de obtención de información menos recomendable y utilizado.

Redundancia y control de consistencia. Examinemos ahora las cosas desde el punto de vista de la redundancia de la información que se acarrea en la tarea.

Si tenemos cuatro estímulos (A, B, C, y D) y los presentamos (k=n) juntos a un sujeto para que elija el mayor en la cualidad, el sujeto resuelve la tarea, nos da esta información y no se produce ninguna redundancia.

Observemos que sucede si tenemos cuatro estímulos y los presentamos en tríadas (k=3). Como hemos visto hay cuatro tríadas posibles:

1. ABC
2. ABD
3. ACD
4. BCD

Supongamos que le presentamos al sujeto la primera tríada y que el sujeto escoge el estímulo A. Entonces el sujeto nos ha dado la siguiente información:

A › B

A › C

Con ello no sabemos nada sobre la relación entre B y C.

Supongamos que a continuación le presentamos la segunda tríada: ABD. ¿Puede un sujeto consistente escoger cualquiera de los tres estímulos? Obviamente, un sujeto consistente no puede ahora escoger B, porque ha establecido ya que A › B.

Sólo le quedan dos opciones, escoger D, si D › A; o escoger A, si A› D. En realidad, dada la respuesta a la primera tríada, la segunda se ha convertido en un par a efectos de elección.

Supongamos que elige D, por ejemplo. Esto significa que:

D › A

De donde se deduce, junto a la información ya tomada que:

Dado que A › B entonces D › B

Dado que A › C entonces D › C

con lo que sólo nos queda por saber la relación entre B y C para tener los cuatro estímulos ordenados.

Pero esa información es la que no podemos establecer debido a que utilizamos tríadas. En efecto, la tercera tríada (A,C,D) no puede aportar nada, dado que un sujeto consistente sólo puede informar que D es el mayor, para ser coherente. Y la cuarta tríada (B,C,D) tampoco puede aportar nada, dado que un sujeto consistente sólo puede responder de nuevo D.

Después de la respuesta a la primera tríada, parecía lógico presentar el par (A,D) -en lugar del trío (A,B,D); y después de la respuesta a este par en el sentido de D › A, sólo quedaba por presentar el par (B,C). Esto sugiere que el máximo de información podría conseguirse variando k en función de las respuestas y adaptando las preguntas a las respuestas del sujeto -lo cual sólo es posible si la administración de los estímulos es individual o computerizada; o bien obviando estos problemas mediante métodos ordinales que después veremos.-

Veamos que sucede si k=2 . En ese caso, para cuatro estímulos, tenemos los siguientes grupos de dos:

1. AB	4. BC	6. CD
2. AC	5. BD	
3. AD		

Supongamos que lo que en definitiva piensa el sujeto es:

$$A > B > C > D$$

Entonces los pares con información clave son:

1. AB   4. BC   6. CD

y los otros tres sólo dan información redundante que se desprende de la que ofrecen los pares 1, 4 y 6. Sin embargo, con k=2 no se puede saber a priori qué pares contienen la información clave. Si el sujeto tuviera en mente una ordenación como:

$$C > A > D > B$$

la información de los pares 1, 4, y 6 únicamente resultaría confusa. Pero en ese caso la de los pares:

2. AC   3. AD   5. BD

sería necesaria y suficiente, siendo la de los otros tres pares redundante.

En general, para cualquier valor de n, con k=2, sólo son necesarios y suficientes n-1 pares, de modo que la información producida por los

$$\binom{n}{2} - (n-1) = \frac{n^2 - n}{2} - (n-1) = \frac{n^2 - 3n + 2}{2}$$

pares restantes es redundante.

La información producida por $\binom{n}{2}$ pares viene expresada por n-1.

Es decir, obsérvese que al aumentar el número de estímulos n a escalar, el número de pares crece muy aprisa sin embargo, el número de pares relevantes crece despacio.

Dicho de otra forma, al aumentar el valor de n el número de pares redundantes crece muy aprisa, mientras que el número de pares significativos crece despacio. La diferencia entre información significativa y redundante crece muy rápidamente en función de n.

Para n=2 tenemos un par significativo y cero redundantes. Para n =3 tenemos dos pares significativos y uno redundante. Si n=4 entonces tendremos tres pares significativos y tres redundantes. Si n=5 tendremos cuatro significativos y seis redundantes. Si n=10, de los 45 pares que se pueden formar sólo 9 serán significativos y los 36 restantes redundantes.

Sea cual sea el valor de k, toda la información redundante puede utilizarse con el propósito de controlar la consistencia de los sujetos en sus respuestas. Una alta inconsistencia a veces es interpretada en el sentido de que los estímulos no son escalables en una única dimensión o de que no son escalables en absoluto.

En cualquier caso, el investigador o el profesional no necesitan normalmente grados tan altos de redundancia, lo que suele llevar a preferir métodos de captación de información más sencillos y ligeros.

Paradoja entre consistencia de los jueces y precisión del escalamiento. Si todos los sujetos están de acuerdo en que un estímulo concreto A es el que posee en más grado una cualidad, entonces en el caso n=k, todos elegirán ese estímulo, y su puntuación será N. Ello implica, como acabamos de ver, una situación de escalamiento muy pobre: el estímulo A es el mayor y no sabemos nada acerca de todos los demás estímulos, excepto que son menores que A en la cualidad bajo consideración. Sin embargo, desde la perspectiva de la consistencia entre los jueces, ello indica una perfecta consistencia. Es decir, si todos los jueces eligen unánimemente un mismo estímulo como el mayor, entonces los jueces actúan consistentemente unos con otros.

Imaginemos la situación contraria, que la muestra de jueces se distribuye entre los diversos estímulos. Por ejemplo, supongamos que tenemos 100 jueces (N=100), y cuatro estímulos (n=4), presentados en un único grupo a cada juez (k=4). Supongamos que de los 100 jueces, 40 dicen que A es el mayor, 30 que lo es B, 20 que lo es C, y 10 que lo es D. Parece obvio que:

$$A > B > C > D ,$$

en función del número de elecciones que reciben como estímulo mayor en la cualidad bajo consideración. En una situación así tendremos información para escalar graduadamente, de un modo preciso, cada ítem; pero, sorprendentemente, ¡ello es así debido a que los jueces son inconsistentes entre sí!

Si N personas juzgando un mismo conjunto de estímulos, juzgan realmente los estímulos por sí mismos, siendo los estímulos estables en sus cualidades y los jueces independientes entre sí, cabría esperar que fueran consistentes en sus juicios. Cuando podemos escalar gradualmente, de un modo preciso, los estímulos, lo estamos haciendo a costa de la inconsistencia de los jueces. Es decir, en la medida en que tenemos precisión ello denota inconsistencia, y en la medida en que tenemos inconsistencia de ella derivamos precisión.

Esta situación puede verse como una versión interesante de la paradoja entre consistencia y precisión. Dos términos que la teoría clásica de tests ha sido proclive a confundir en una sola cosa y que, en realidad, pueden ser, paradójicamente, contradictorios.

Este razonamiento es patente en el caso particular k=n , pero ¿qué sucede en el caso k=n-1, k=n-2, etc.? El razonamiento puede extenderse. En la medida en que dos jueces difieren en un juicio sobre los mismos estímulos son inconsistentes entre sí -lo que contribuye a ponerlos en duda como jueces y a poner en duda el procedimiento de jueces como medio para escalar-. Por otro lado, si los N jueces fueran completamente consistentes entre sí ¿no bastaría con un solo juez? Si dos jueces discrepan es posible que sea debido a que tienen criterios de juicio diferentes o a algún tipo de error aleatorio. En cualquier caso ello merece investigación. No pueden tomarse sin más las diferencias entre jueces como información que aporta un escalamiento más fino.

La información métrica como producto de la inconsistencia. El problema puede verse de otro modo. Al presentar estímulos a un sujeto y pedir que escoja uno de entre ellos en función de una cualidad, el sujeto sólo nos brinda información ordinal. Nos dice el orden de desigualdad de los estímulos. Ya hemos visto que los grados de desigualdad que pueden extraerse es una función de la relación entre n y k. Al utilizar N jueces sobre los n estímulos, si los N jueces fueran consistentes consigo mismos y entre ellos, arrojarían unánimemente el mismo orden de desigualdad, juicio a juicio, sin excepción; y el investigador sólo podría establecer el orden de desigualdad entre los n estímulos en n-k+1 grados, independientemente del N de la muestra. Con jueces perfectamente consistentes daría igual tener N=1 que N=1000, la única información que se obtendría serían los n-k+1 grados de desigualdad entre n estímulos. Si los jueces son consistentes, sólo podemos saber que A›B›C..., sin que podamos saber nada acerca del grado en que A es mayor que B en la cualidad o del grado en que B es mayor que C en la cualidad.

Sin embargo, cuando se usan N jueces aparecen discrepancias en la información ordinal que dan; es decir, aparecen inconsistencias en sus juicios de desigualdad. Esas inconsistencias se traducen en las frecuencias con que un estímulo es juzgado mayor que otro.

Por ejemplo, en un par (A,B) el 80% de los jueces puede afirmar A›B y el 20% restante B›A, mientras que en un par (B,C) el 60% de los jueces dice B›C mientras que el 40% restante dice C›B. Algunos métodos de escalamiento utilizan de diferentes modos esas frecuencias o transformaciones de las mismas para calcular las distancias entre los estímulos.

Estos métodos razonan más o menos así: En el caso (A,B) si el 80% dice A›B y sólo el 20% dice B›A debe ser que la diferencia entre A y B es más clara, y *por tanto mayor*, que la diferencia entre B y C. El método concluye por fin que A›B›C, pero además, que la distancia entre A y B es mayor que la distancia entre B y C, y todavía más, cuantifica de algún modo esa distancia a partir de las diferencias en frecuencias, por ejemplo:

Razonamiento:	Valores escalares asignados:
Origen arbitrario de la escala = 0.	Valor escalar de C = 0
Distancia C a B = 60-40 =20:	Valor escalar de B = 0 + 20 =20
Distancia A a B = 80-20=60	Valor escalar de A = 20 + 60 = 80

Escala resultante:

0	10	20	30	40	50	60	70	80	90	100
C		B						A		

Muchos métodos, como veremos, introducen un tratamiento más sofisticado de esos porcentajes, pero esencialmente éste es el tipo de resultado.

La paradoja, (y el 'mérito' de estos métodos,) consiste en que convierten las discrepancias entre jueces, una información sobre la inconsistencia de los mismos, en dato para la obtención de información cuantitativa sobre las distancias entre estímulos. Convierten la inconsistencia en fuente de precisión. Por eso puede afirmarse que la información ordinal que arrojan estos métodos, basados en tareas de elección, es fruto de la discriminación entre los objetos que ofrecen los jueces, pero la información métrica es fruto de la inconsistencia de los jueces.

Desde este punto de vista queda claro que lo que se está escalando no son solamente los estímulos en función de sus cualidades, sino también la inconsistencia de los jueces. Esa inconsistencia se acepta a veces bajo el argumento de que lo que en realidad se escalan son, en todo caso percepciones humanas; algo que ya comentamos al hablar del carácter subjetivo del escalamiento.

Relación entre los métodos de estímulo simple, de elección de alternativas, de comparación de estímulos por pares y de comparación de estímulos por grupos. En primer lugar, tiene que resultarle ahora obvio al lector que el método de comparación de estímulos por pares es tan solo un caso particular, con k=2, del método de comparación de estímulos por grupos. Simplemente, la comparación de estímulos por pares es un método especialmente popular que permite obtener la máxima información posible en comparación de estímulos. Los principios e indicaciones expresados para un método sirven, adaptados, para el otro.

En segundo lugar, debe quedar claro que ambos, la comparación de estímulos por pares y por grupos, son métodos de elección. La tarea del sujeto consiste en elegir un estímulo, el que más posea de una propiedad (orientación de juicio) o el que más se adecue a su experiencia, actitud, etc (si la orientación es de respuesta).

¿Qué sucede si se efectúa un método de este tipo con k=1? Este caso puede interpretarse como un absurdo, dado que el sujeto no tiene elección, y por tanto, desconsiderarse. Pero también es posible reinterpretar el método del estímulo simple como un método de elección con k=1. En este caso, los estímulos son presentados uno a uno y el sujeto decide si cada estímulo tiene o no, en suficiente grado, una cualidad.

Desde el método de presentar los estímulos uno a uno (k=1) hasta el método de presentarlos todos a la vez (k=n) todos pueden ser considerados métodos de elección.

A su vez, el método de elección de alternativas puede considerarse una variante del método de estímulo simple donde la elección del estímulo se presenta graduada. Con este punto de vista, los cuatro métodos vistos hasta aquí (estímulo simple, elección de alternativas, comparación por pares y comparación por grupos) pueden considerarse métodos de elección.

5 . 5

5. *Ordenación de estímulos por grupos.*

En una tarea de ordenación de estímulos por grupos se forman todos los grupos o combinaciones posibles de k elementos y se le muestran una a una al sujeto. El sujeto debe ordenar los estímulos o enunciados en cada grupo de mayor a menor (o de menor a mayor) en la dimensión considerada.

Si tenemos n elementos a escalar utilizando un procedimiento de ordenación para captar información de ordenación de estímulos por grupos de k, entonces el valor de k determina exactamente el tipo de tarea. Cuando hablamos de ordenación de estímulos por grupos nos referimos a la situación en que:

$$3 \leq k \leq n$$

Casos que pueden presentarse en función del valor de k. El caso con $k = n$ es de uso frecuente y dadas sus particularidades lo estudiaremos aparte, como el sexto método de obtención de información, aunque como se ve, no es más que un caso particular de este método.

Si k=1 entonces la tarea se reduce a presentar un solo estímulo cada vez. La tarea de ordenar un solo estímulo respecto a una dimensión puede considerarse sin sentido y desconsiderarse; pero también es posible reinterpretar este caso en el sentido de considerar si un estímulo ocupa una posición ordinal tan alta como para considerar que presenta la cualidad. Bajo esta reinterpretación k=1 se reduce a una tarea de estímulo simple.

Si k=2 entonces los estímulos son presentados en grupos de dos y se solicita al sujeto que diga cuál es el mayor y el menor en la cualidad. Evidentemente, decir cuál es el mayor y cuál el menor en una cualidad entre dos estímulos equivale a solicitar que se escoja el mayor (el otro necesariamente será el menor de los dos). De este modo queda claro que el método de ordenación de estímulos por grupos de dos equivale a una tarea de comparación de estímulos por pares.

Ejemplo. Veamos un ejemplo de tarea de ordenación de estímulos por grupos. Sean los enunciados sobre política social:

A. La Administración Pública debe favorecer el bienestar público mediante un reparto equitativo de la riqueza a través de impuestos directos.

B. Las fuertes desigualdades en riqueza que produce el mercado son positivas, actuando como estímulo para la superación individual.

C. Los servicios públicos y sociales deben suprimirse o reducirse a la mínima expresión, para bajar los impuestos directos al máximo.

D. Los servicios sociales, seguridad social, pensiones etc. son prioritarios, aunque deban sostenerse sobre un impuesto sobre la renta elevado.

Queremos escalarlos en función de la propiedad "representar una política liberal de mercado".

Con n=4 y k=3 podemos obtener 4 combinaciones o grupos:

$$\binom{n}{k} = \binom{4}{3} = \frac{4!}{3!(4-3)!} = 4$$

Obtenemos esos cuatro grupos de tres elementos:

1. ABC
2. ABD
3. ACD
4. BCD

Antes de comenzar a presentar las ternas se le dan al sujeto las siguientes instrucciones:

> Los enunciados que se le van a presentar a continuación en grupos de tres representan diversas orientaciones políticas más o menos propias de una política liberal de mercado. Para facilitar la tarea le vamos a presentar los enunciados en grupos de tres. Ordene los tres enunciados de cada grupo de más a menos representativo de una política de "liberalismo económico". Observe que no le solicitamos que nos diga si está Ud. o no de acuerdo con los enunciados; independientemente de cuál sea su opinión sobre el tema, ordénelos en función de cuán propios son de una política liberal de mercado. Para ello, en cada trío de enunciados, escriba un 3 al lado del enunciado de ese trío más propio de un liberalismo económico, un 1 al lado del menos propio del liberalismo económico y un 2 al enunciado restante.

A continuación se le van presentando las ternas, una a una, y el sujeto va calificando cada estímulo en cada terna con un número del 1 al 3.

Puntuación máxima y mínima. En cada tríada, cada enunciado puede haber recibido un 1 (posición más baja), un 2 (intermedia) o un 3 (posición más alta). Obsérvese que, en este sistema de puntuación, la puntuación máxima de cada tríada depende de k. Es decir, en cada grupo los estímulos se puntúan de 1 hasta k; en este caso de 1 hasta 3.

Podemos determinar la puntuación total de cada uno de los estímulos, para un sujeto, como la suma de puntos que ha obtenido a través de todos los grupos a los que pertenece.

En este caso, cada elemento está en 3 de las cuatro tríadas:

$$\binom{n-1}{k-1} = \binom{3}{2} = \frac{3!}{2!(3-2)!} = 3$$

Por tanto la puntuación máxima que un estímulo puede obtener es:

$$k \cdot \binom{n-1}{k-1} = 3 \quad 3 = 9$$

y la mínima:

$$1 \cdot \binom{n-1}{k-1} = 1 \cdot 3 = 3$$

Así que en este caso concreto, con n=4 y k=3 la escala va desde 9 (máximo), hasta 3 (mínimo). Obsérvese que, dado el sistema de puntuación, la escala hace mínimo en:

$$\binom{n-1}{k-1}$$

En general, la puntuación máxima que un estímulo puede obtener de un sujeto en este método de obtención de información es:

$$k \cdot \binom{n-1}{k-1} = \frac{k\,(n-1)!}{(k-1)!\,(n-k)!}$$

y la mínima, como acabamos de ver:

$$\binom{n-1}{k-1} = \frac{(n-1)!}{(k-1)!\,(n-k)!}$$

A veces puede resultar más simple mantener la escala de puntos con el mínimo en 0. En ese caso basta con restar

$$\binom{n-1}{k-1}$$

a cada puntuación, quedando el máximo en:

$$k \cdot \binom{n-1}{k-1} - \binom{n-1}{k-1} = \frac{k\,(n-1)!}{(k-1)!\,(n-k)!} - \frac{(n-1)!}{(k-1)!\,(n-k)!}$$

lo que puede escribirse de modo más simple como:

$$(k-1) \cdot \binom{n-1}{k-1} = \frac{(k-1)\,(n-1)!}{(k-1)!\,(n-k)!} = \frac{(n-1)!}{(k-2)!\,(n-k)!}$$

y el mínimo en:

$$\binom{n-1}{k-1} - \binom{n-1}{k-1} = 0$$

Otro modo de obtener el mismo resultado es puntuar cada estímulo en cada grupo desde 0 hasta k-1.

Si tenemos N sujetos a los que hemos enfrentado con la tarea, entonces, para una escala de puntuación con origen en 0, el máximo para los N sujetos será:

$$N \cdot (k-1) \cdot \binom{n-1}{k-1} = N \cdot \frac{(k-1)\,(n-1)!}{(k-1)!\,(n-k)!} = \frac{N \cdot (n-1)!}{(k-2)!\,(n-k)!}$$

y el mínimo permanecerá en 0.

Una vez que todos los estímulos han recibido una puntuación total entre 0 y el máximo que acabamos de ver, es fácil ver cómo estos puntos podrían utilizarse como valores de escala o bien transformarse de algún modo en la escala final, (por ejemplo haciéndolos proporcionales a 100).

Redundancia y control de consistencia. Examinemos ahora la cuestión desde otro punto de vista. Como señalamos, volviendo a nuestro ejemplo, con cuatro enunciados se obtienen cuatro ternas.

1. ABC
2. ABD
3. ACD
4. BCD

Ahora bien ¿es necesario utilizar tantos grupos en una tarea de ordenación? ¿Cuánta información redundante hay? Veamos que sucede cuando hipotéticamente un sujeto contesta cada una de estas ternas.

Primero al sujeto se le presenta la primera terna:

A. La Administración Pública debe favorecer el bienestar público mediante un reparto equitativo de la riqueza a través de impuestos directos.

B. Las fuertes desigualdades en riqueza que produce el mercado son positivas, actuando como estímulo para la superación individual.

C. Los servicios públicos y sociales deben suprimirse o reducirse a la mínima expresión, para bajar los impuestos directos al máximo.

Y supongamos que el sujeto la puntúa del siguiente modo:

C=3; B=2; y A=1

Esto significa que ha establecido dos grados de desigualdad:

C › B › A

Obsérvese que ¡sólo nos queda saber que posición ocupa el estímulo D para tener la ordenación completa!

Supongamos que la siguiente terna que se presenta es la ABD, y que, teniendo en cuenta el contenido de D:

> D. Los servicios sociales, seguridad social, pensiones etc. son prioritarios, aunque deban sostenerse sobre un impuesto sobre la renta elevado.

el sujeto puntúa de este modo:

D=1; A=2 ; y B= 3.

Es decir, el sujeto ha ordenado los items de la terna así:

B › A › D

Podemos observar que:

1. Ya tenemos la escala completa ordenada:

   C › B › A › D

   por tanto ya no necesitamos presentar ninguna terna más; toda la información que nos dieran las otras dos ternas sólo serviría como control de consistencia del sujeto.

2. Aun dentro de esta terna, la información:

   B › A

   es redundante, sólo sirve como control de consistencia.

Si la elección del sujeto hubiera sido de otro modo también se habría producido un alto nivel de redundancia. Salvo para propósitos de investigación de cómo operan los métodos de escalamiento, los investigadores buscan los procedimientos efectivos más sencillos y menos costosos. La mayoría de las veces el investigador no necesita tanta redundancia de información como control de consistencia. Por eso, el procedimiento de captación de información mediante ordenación de estímulos por grupos no resulta práctico.

El procedimiento para obtener información mediante ordenación sobre un conjunto de estímulos reduciendo al máximo la redundancia consiste en presentar todos los estímulos a la vez, caso que estudiamos a continuación.

5.6

6. Ordenación del grupo total de estímulos.

Este es el caso de ordenación con n=k. Simplemente se presentan al sujeto una sola vez todos los estímulos a escalar y se le solicita que los ordene de menos a más (o de más a menos) en la dimensión o cualidad considerada.

Para hacerlo se le puede pedir que de el valor 1 al menor , 2 al siguiente, etc. Al valor mayor le dará en este caso el valor n. Con este procedimiento cada estímulo obtiene para cada sujeto una puntuación entre 1 y n. Para N sujetos, por tanto, entre N y Nn. A veces se considera más fácil manejar una escala de puntos entre 0 y n-1; con lo que en una muestra de N sujetos cada estímulo puede haber recibido una puntuación total entre 0 y N(n-1).

Cuando el investigador o el profesional optan por una tarea de ordenación de estímulos suelen preferir esta variante, con n=k, lo que da una tarea clara, que los sujetos entienden y aceptan bien y que resulta relativamente sencilla y rápida, permitiendo obtener el orden completo de los estímulos.

Carácter ipsativo de los datos de ordenación de estímulos. Los datos obtenidos de este modo tienen patentemente una característica: el rango de puntuaciones que un sujeto puede escoger para un estímulo depende del rango de puntuaciones que haya fijado previamente para el/los estímulos anteriormente enjuiciados. Es decir, la puntuación de unos estímulos depende de la que han recibido otros. En el caso del último estímulo a juzgar la dependencia es total, por eso para describir n estímulos basta con conocer n-1 puntuaciones.

Veámoslo con un ejemplo. Supongamos que queremos que un sujeto ordene cuatro estímulos A,B,C y D. Para simplificar, supongamos que la estrategia del sujeto es localizar el mayor, asignarle la puntuación mayor (4), localizar el siguiente mayor, asignarle la puntuación siguiente, etc. El sujeto puede por ejemplo, identificar A como el mayor y por tanto A=4. Para los siguientes 3 estímulos ahora sólo dispone de tres puntuaciones. Supongamos que identifica D como el siguiente mayor, D=3. Ahora sólo le quedan dos estímulos a los que sólo puede asignar los valores 1 ó 2. Supongamos que identifica C=2. Por último, la puntuación de B depende completamente de la que han obtenido los otros tres estímulos. Desde luego, si el sujeto sigue la estrategia de asignar su valor a cada estímulo según se le presentan -o cualquier otra- se produce la misma dependencia. Inevitablemente con este método una puntuación depende completamente de las otras asignadas; las puntuaciones asignadas a los estímulos no son juicios independientes entre sí.

La cuestión de la dependencia tiene su importancia porque normalmente se aspira a que los datos obtenidos puedan ser objeto de diversos tratamientos estadísticos, según las necesidades del profesional o el investigador, y resulta que muchas técnicas estadísticas parten del supuesto de puntuaciones independientes. La dependencia entre observaciones -sobre todo si es fuerte o extrema, como en este tipo de datos- crea dificultades a una buena parte de las técnicas estadísticas.

Paradoja de la dependencia entre puntuaciones. No obstante, es paradójico que se reconozca claramente esta ipsatividad a las puntuaciones obtenidas de ordenaciones y no a otras. Consideremos de nuevo los cuatro estímulos. Si un sujeto es consistente y mantiene un patrón determinado, por ejemplo,

$$A > D > C > B$$

entonces, sólo hay realmente tres grados de desigualdad que establecer sea cual sea el método que se utilice para expresar esto. Vistas así las cosas, las respuestas obtenidas por comparación de estímulos por pares o por comparación de estímulos por grupos, *deberían ser* dependientes, en tanto que, como ya mostramos anteriormente, son en buena parte redundantes. Si los jueces son consistentes entre sí, como ya vimos, además, cada juez debería ser completamente redundante con otro, si operan del mismo modo sobre la misma cualidad de los mismos objetos. La no dependencia entre las puntuaciones obtenidas por medio de estos otros métodos aparece como fruto de inconsistencia de los jueces. Lo que diferencia al método de ordenación total de estímulos respecto a estos otros en este sentido es que el método de ordenación total de estímulos fuerza al sujeto a ser consistente consigo mismo.

Dicho de otro modo, con n estímulos, si entre esos n estímulos tiene sentido la relación "ser mayor que" en una dimensión o cualidad, necesariamente sólo se pueden establecer n - 1 grados de desigualdad. Expresado en otros términos, si la operación ">" tiene sentido entre n estímulos, las puntuaciones asignadas a los mismos que reflejen esas relaciones necesariamente son ipsativas y necesariamente con la descripción de n-1 estímulos es suficiente. Ello es tanto como decir que hay necesariamente una posición redundante en la concepción de un modelo consistente de ordenación de estímulos en una dimensión.

El método de ordenación total de estímulos impide materialmente al juez ser inconsistente; le fuerza a expresar un modelo de relaciones de desigualdad consistente. Pero, lo paradójico es que la dependencia no está en el método sino en el modelo de estímulos subyacente.

Cuando un modelo del tipo comparación de estímulos por pares o por grupos, no arroja puntuaciones o respuestas dependientes, ello es fruto de la inconsistencia de los jueces -o, como a veces se dice, de la inescalabilidad de los estímulos en una dimensión.