

PARTE TERCERA
APROXIMACION HISTORICA.

5. ANTECEDENTES DE LA MEDICION, PSICOFISICA Y ESCALAMIENTO.

1. Antecedentes de la medición.

Algunos autores han pretendido encontrar algunos antecedentes remotos de la medición psicológica en un pasado lejano. Por ejemplo, Adams y Schvaneveldt (1.985), en un capítulo titulado "Niveles de medida y escalación", han considerado el pasaje bíblico del libro de los Jueces (7:2,8) en que Jerubaal selecciona (con considerable ayuda de Yavéh) un grupo de trescientos soldados de entre treinta y dos mil hombres, por un procedimiento de eliminación múltiple sucesiva basado en la observación de algunas conductas de los mismos, como un antecedente de aplicación de . A nuestro juicio hay que alejarse mucho de una exégesis razonable para encontrar "tests" en el libro de los Jueces.

Otros, como primer antecedente, han resaltado el hecho de que Hiparco ciento cincuenta años antes de Cristo utilizase una escala de seis puntos para ordenar las estrellas en función de su brillo percibido, lo que puede considerarse el primer ejemplo de utilización de una escala subjetiva para ordenar impresiones. Otros, en fin, mencionan que la medición educativa se ha usado en China desde tiempos remotísimos para la selección de funcionarios. Sea como fuere, las necesidades de medición, en el campo de las ciencias de la conducta, aparecieron prontamente en la historia de la humanidad asociadas a problemas de selección y clasificación de sujetos o a cuestiones de percepción subjetiva de objetos. Dos temas que pueden considerarse claramente representados en los desarrollos posteriores.

Sin embargo, estos antecedentes pueden ser considerados anecdóticos y hay que llegar hasta la segunda mitad del siglo pasado para encontrar los acontecimientos que pueden considerarse más propiamente fundacionales de la psicometría y también de la psicología. Como es sabido, frecuentemente se ha tomado la creación de un laboratorio de psicología en Leipzig en 1.879 por Wilhelm Wundt como fecha fundacional de la psicología. Sin embargo, la orientación estructuralista de Wundt y su escasa atención a la cuestión de las diferencias individuales lo hacen poco propicio para ser considerado en el origen de la psicometría. En su lugar, las personas y obras de Fechner, dando cuerpo a la psicofísica, y de Galton y J.M. Cattell, en el origen de los primeros tests de tipo sensorial y motor, constituyen candidatos más idóneos

para la fundación de los dos grandes troncos de desarrollo de la psicometría: el escalamiento y la teoría de los tests. Desde una consideración psicométrica ambas aportaciones pioneras tienen un considerable valor histórico.

2. El origen del escalamiento psicofísico.

Fue Gustav Theodor Fechner (1.801-1.887) quién definió la psicofísica como una ciencia exacta acerca de las relaciones funcionales de dependencia entre el cuerpo y la mente. Para Fechner la psicofísica incluía tanto la medida de los atributos sensoriales como la cuantificación de la percepción de los sujetos, de forma que se pudieran poner en relación los continuos psicológicos y los continuos físicos de los estímulos.

Gustav Theodor Fechner fue, además de físico, y sobre las demás cosas un filósofo con una mente abierta y curiosa. Publicó hasta siete obras de filosofía entre 1.836 y 1.879, que él consideraba al parecer sus principales obras, y, entre otras curiosidades, dedicó cuatro años al estudio de las post-imágenes por el método de fijar la vista durante prolongados periodos sobre el sol. Desde el punto de vista de este hombre, la psicofísica solo fue un trabajo menor y práctico con el propósito fundamental de mostrar que sus ideas filosóficas tenían una utilidad. El momento de surgimiento de la psicofísica está bien documentado gracias al propio Fechner.

Al parecer fue en la tarde del 22 de Octubre de 1.850, cuando Gustav Theodor Fechner formuló en su mente la hipótesis de que a incrementos geométricos de la intensidad de los estímulos corresponderían incrementos aritméticos de la sensación de los sujetos. Según parece Fechner pasó la mañana de aquel día en la cama, meditando sobre la necesidad de medir la 'energía mental' en correspondencia a como se efectuaban mediciones de la energía en física. Una primera expresión de sus pensamientos acerca de la energía mental apareció en una obra titulada "Sobre la naturaleza de los cielos y el mundo no material" que publicó en 1.851. Pero la idea acerca de la relación matemática entre estímulos y sensaciones, fue plasmada en la llamada ley de Fechner, publicada en sus "Elemente der Psychophysik" de 1.860, y fue aceptada, en general, durante los cien años siguientes, soportada por datos relativos a la discriminabilidad de los estímulos. En realidad, Fechner se había beneficiado de una intuición de Weber formulada vagamente en torno a 1.800. Según Weber proporciones iguales de estímulo deberían producir diferencias iguales en el juicio. Esta idea fue formulada

matemáticamente por Fechner, y acuñada por este último como la ley de Weber.

Después de su trabajo de 1.860, Fechner nunca mostró demasiado interés por la psicofísica, limitándose a hacer una replica a las críticas recibidas en su "Sachen der Psychophysik" (Sobre cuestiones psicofísicas) de 1.877 y a una revisión de sus "Elemente" que publicó en 1.882 bajo el título de "Revision der Hauptpunkte der Psychophysik" (Revisión de puntos centrales de la Psicofísica). En esos años Fechner estaba más interesado en la cuestión de cómo medir la belleza, y algunas de sus intuiciones sobre este tema pueden situarse también en los orígenes de las técnicas de escalamiento.

Sin embargo, Fechner si que tuvo clara conciencia de haber dejado una orientación válida que perduraría durante los años siguientes. Stevens en su artículo "On the psychophysical law" de 1.957 (Psychological Bulletin, Vol. 64) citó unas premonitorias palabras de Stevens: "La Torre de Babel nunca fue terminada porque los trabajadores no llegaron a ponerse de acuerdo sobre cómo construirla; mi edificio psicofísico se mantendrá en pie porque los trabajadores nunca estarán de acuerdo sobre cómo derribarlo."

Precisamente, la ley logarítmica de Fechner sería alterada en los años cincuenta cuando S.S. Stevens propuso una ley potencial que se basaba en datos de estimación de magnitudes. Desde ese momento se produjo un resurgimiento de la investigación psicofísica con una serie de trabajos que ponían a prueba y comparaban ambas aproximaciones. Helm, Messick y Tucker (1.961) obtuvieron una relación exponencial entre las escalas de estimación de magnitudes y las escalas de discriminación. Galanter y Messick (1.961) confirmaron esa relación. La importancia de estos trabajos estriba en que permitían considerar como ciertas simultáneamente las leyes de Fechner y Stevens. Las escalas de estimación de magnitudes, como Stevens proponía, podían mantener una relación potencial con el estímulo, mientras que las escalas de discriminación, como mostró Fechner, mantendrían una relación logarítmica con la magnitud estimular. Una serie de trabajos, históricamente anteriores a los que acabamos de mencionar, se habían ocupado de establecer los métodos psicofísicos de trabajo, tales como el método de ajustes, de los límites, de los estímulos constantes, etc. El mismo Thurstone, de quién después nos ocuparemos con más detalle, había realizado importantes contribuciones a la psicofísica, y se produjeron también una serie de extensiones de los métodos de

Fechner y de Thurstone, como el modelo de Bradley-Terry-Luce y el de Dawkis (Baird y Noma, 1.978).

En 1.960 la Psychometric Society pudo celebrar simultáneamente el veinticinco aniversario de su fundación y el primer centenario de la publicación de los "Elementos de Psicofísica". Se organizaron una serie de sesiones especiales cuyos principales trabajos fueron recogidos después en Psychometrika en 1.961. Entre esas sesiones especiales destaca la titulada "Psicofísica: Cien años después" con importantes trabajos de Boring (1.961), acerca del marco fundacional de la psicofísica, de Gulliksen (1.961), con una revisión sobre el escalamiento lineal y multidimensional, y de Stevens (1.961), Smith (1.961) y Swets (1.961). En otra sesión titulada "Nuevos desarrollos en la psicología matemática" se presentaron trabajos de Estes (1.961), Green (1.961) Gulliksen (1.961) y Guilford (1.961).

Importancia de la psicofísica. La importancia de la psicofísica en psicología se debe a una serie de razones principales. Por una parte, a su carácter fundacional en el uso de métodos experimentales en psicología, con una formulación matemática de sus principios. Por otra, a lo sustantivo de la temática de que se ocupa, enlazando lo físico y lo psíquico, con todos los aspectos esenciales de esta temática tan del gusto del fundador del campo. Pero, no solo son razones intrínsecas de método y contenido las que hacen importante a la psicofísica. Además, la psicofísica ha jugado un papel ejemplar para otras disciplinas y campos de la psicología y sus métodos se han extendido y adaptado para utilizarse sobre otros contenidos psicológicos con o sin referencia a continuos físicos reales.

En tiempos recientes los avances y métodos psicofísicos han saltado de nuevo a otros campos de aplicación tales como cuestiones de psicología clínica (Grossberg y Grant, 1.978) y cuestiones de psicología social (Wegener, 1.982). Pero sin duda, la gran extensión y fruto de la psicofísica fuera de si misma, y la aportación de un mayor interés psicométrico, ha sido el dar origen a las técnicas de escalamiento psicológico de un modo más o menos directo.

Tareas del escalamiento psicofísico. El problema fundamental de la psicofísica es la medida de la magnitud de la sensación. Medida que no puede efectuarse de un modo directo y que debe ser fruto de inferencia a partir de las respuestas observables de los sujetos a estímulos.

La psicofísica tuvo que enfrentar el problema del escalamiento de las sensaciones de los sujetos. Esta tarea implica

que el sujeto produzca respuestas sensoriales a estímulos. Siguiendo a Geschneider (1.988) esas respuestas pueden ser divididas en tres tipos:

"a) Aquellas que requieren que el sujeto haga juicios de discriminación ordinal simple de estímulos (e.g. Fechner 1.860, Thurstone 1.927),

b) Aquellas que requieren que los sujetos ajusten estímulos a la partición del continuo sensorial en intervalos subjetivamente iguales (e.g. bisección o equisección), y

c) Aquellas que requieren de los sujetos que asignen números a estímulos que presumiblemente representan magnitudes de sensación (e.g. estimación de magnitudes o escalamiento categórico) o ajustar estímulos para igualarlos a números presentados por el experimentador (producción de magnitudes)." (Gescheider, 1.988).

Las respuestas en la tarea de escalamiento son una función conjunta de factores cognitivos y sensoriales. Estas tareas para el escalamiento de sensaciones nacen en el campo de la psicofísica, utilizadas para obtener respuestas ante estímulos sensoriales, y después serán trasladadas al campo del escalamiento psicológico, donde se utilizarán para obtener respuestas de los sujetos ante estímulos cuyo referente sensorial no es inmediato. Si se comparan las tareas que cita Geschneider con las que implican los métodos de juicio y de respuesta (Figura 4) podrá apreciarse la relación.

Por supuesto la psicofísica constituye un campo de investigación con contenido y desarrollos propios bien diferenciado de la psicometría. Por ello no entraremos en una consideración más detenida de sus contenidos, avances y características. Sin embargo, desde un punto de vista psicométrico, la psicofísica constituye el origen histórico de los procedimientos de escalamiento. *El escalamiento fue primero escalamiento psicofísico y después, psicológico, y desde ese punto de vista la fundación de la psicofísica puede considerarse también la fundación del escalamiento.*

Es desde esta orientación de la relación entre psicofísica y escalamiento que nos resultan de interés las aportaciones de la primera. Algunos de los principales trabajos sobre la psicofísica y el escalamiento son el de Guilford "Psychometric Methods" (N.Y. McGraw) de 1.954, el de Edwards

"Techniques of attitude scale construction" (N.Y. Appleton) de 1.957, el capítulo de Galanter "Contemporary Psychophysics" (en Brow y otros, "New Directions in Psychology". N.Y. Holt) de 1.962, y muy especialmente en el libro de Torgerson "Theory and Methods of Scaling" de 1.958. Con una orientación docente destaca el breve libro de Sidney A. Manning y Edward H. Rosenstock "Classical Psychophysics and Scaling" de 1.968, que adopta forma de texto programado, y del que existe una versión castellana, editada por primera vez en 1.971 por Trillas, que ha tenido varias reimpresiones.

George A. Gescheider (1.988) ha presentado una revisión del campo del escalamiento psicofísico muy orientada hacia los trabajos de psicólogos experimentales interesados en los aspectos sensoriales y cognitivos, dejando a un lado el trabajo efectuado por los teóricos de la medida en este campo, justamente la porción de más interés psicométrico. Las contribuciones de la psicofísica relacionadas con el escalamiento pueden encontrarse más adecuadamente representadas en el trabajo de revisión de Cliff (1.973) acerca del escalamiento, así como en la revisión de Carroll y Arabie (1.980) sobre escalamiento multidimensional, y en la de Young (1.984) también dedicada al escalamiento, en continuación de la de Cliff. Sin embargo, la revisión de Cliff prestó también atención a las cuestiones más propiamente psicofísicas, a diferencia de la de Young y a semejanza de la de Gescheider.

El mismo día de la muerte de Fechner nacía L.L. Thurstone, figura pionera en los desarrollos del escalamiento psicológico a la que está estrechamente ligado el tránsito desde la psicofísica al escalamiento psicológico.

3. El tránsito de la psicofísica al escalamiento psicológico.

El énfasis del trabajo de Fechner estuvo puesto en la determinación de la relación entre el estímulo y la sensación, más que en la medición psicológica, pero esta línea de trabajo abrió el camino a las aportaciones de Thurstone. Louis Leon Thurstone (1.887-1.955) es la figura clave en el salto entre la psicofísica y la escalación o escalamiento psicológico (Gulliksen, 1.958). *Thurstone puso de manifiesto que los métodos psicofísicos de escalamiento podrían también utilizarse para obtener escalas de medida acerca de dimensiones y atributos psicológicos carentes, al menos de un modo directo, de una dimensión estimular física manifiesta con la que ponerlos en relación.*

Aunque se ha intentado fijar este salto en algún trabajo concreto de Louis L. Thurstone, quizás sea más indicado admitir que esta nueva aproximación surgió de un modo gradual a lo largo de un conjunto de trabajos de este autor fechados entre 1.925 y 1.930. En 1.925 Thurstone publicó un primer trabajo acerca de un método de escalación de tests psicológicos y educacionales. En 1.926 publica un trabajo acerca del concepto de edad mental y otro sobre la puntuación del desempeño individual. Los dos años siguientes, el 1.927 y el 1.928 suelen considerarse claves en este tránsito hacia la escalación psicológica. En el primero publica 'Tres leyes psicofísicas' y 'Análisis psicofísico' entre otros trabajos de corte psicofísico, pero también 'Una ley del juicio comparativo' y el 'Método de comparaciones apareadas para valores sociales', además de 'Una unidad de medida mental' y 'Una unidad de medida en escalas educativas', entre otros trabajos. El año 28 puede recordarse en la historia de la psicometría por 'Las actitudes pueden medirse' y 'La medida de la opinión', dos trabajos que abrían definitivamente el desde entonces dilatadísimo campo de la medición de las actitudes.

Miller (1.968) ha tratado este tema del nacimiento del escalamiento psicológico desde la psicofísica, bajo la supervisión de Boring, con su sugestivo estilo. Al parecer L.L. Thurstone, que ejercía durante los últimos años de la década de los veinte como profesor de psicofísica en la Universidad de Chicago, pudo llegar a considerar irrelevante el objeto al que se aplicaban tradicionalmente los métodos psicofísicos, "se aburrió de lo que estaba enseñando", en palabras de Miller, y decidió comenzar a usar estos métodos para indagar variables psicológicas que podía considerar más relevantes. clase de lógica psicofísica, está claro que tendríamos la posibilidad de una descripción objetiva de mayor significación psicológica que el limen sensorial.» (Thurstone, 1.959). Las palabras de Thurstone son suficientemente expresivas acerca de cuan conscientemente se produjo este proceso y, a la vez, del modo en que quedan enraizados histórica y conceptualmente psicofísica y escalamiento.

Había sido el mismo Fechner el que había desarrollado el método de las comparaciones apareadas, pero fue Thurstone quién desarrolló la ley del juicio comparativo para este tipo de datos, poniendo de manifiesto que era aplicable a dimensiones psicológicas, obteniendo escalas de medida no relacionadas con un conjunto de estímulos físicos. El *método de las comparaciones apareadas* es un método de escalamiento en el que al sujeto se le presentan todos los pares posibles formados por los estímulos a escalar, y se le solicita que tome de cada par aquel que "tiene más" (o el que "tiene menos") de una característica

determinada. La información sobre las elecciones realizadas por una muestra de sujetos permiten, utilizando un conjunto de ecuaciones conocidas como la "ley del juicio comparativo", escalar los estímulos. Es decir, permiten asignar a cada estímulo un valor de escala, ubicarlo sobre un continuo unidimensional. La ley del juicio comparativo no es otra cosa que un conjunto de ecuaciones mediante las cuales datos acerca de comparaciones de estímulos pueden dar lugar a la posición en una escala común de los estímulos bajo determinados conjuntos de supuestos. Es decir, un conjunto de ecuaciones capaces de escalar objetos de cualquier tipo comparados por los sujetos en alguna propiedad o característica. En los años siguientes los métodos de escalamiento se extendieron a una variedad de campos psicológicos incluyendo aprendizaje, personalidad, estética, relaciones interpersonales etc. Gulliksen (1.958) habla de "varios cientos de artículos relevantes" en diversos campos biométricos, psicológicos, de la agricultura, la óptica etc.

Debe mencionarse que Louis Leon Thurstone, como es sabido, contribuyó a la psicometría no solo desde sus trabajos sobre escalamiento sino también con sus aportaciones en el campo del *análisis factorial*, enraizados en el desarrollo de los modelos conceptuales sobre la inteligencia y las aptitudes primarias (1.938; 1.941). Esos trabajos se hallan bien representados en su obra de 1.947 'Análisis Factorial Múltiple'. El análisis factorial, en sentido amplio, es un conjunto de técnicas estadísticas multivariadas que permiten analizar la "estructura" de relaciones entre un conjunto de variables. Las técnicas de análisis factorial permiten extraer a partir de un conjunto de variables medidas, observadas, un conjunto -menor- de variables latentes que tratan de explicar al máximo la variabilidad de las variables observadas. Quizás el ejemplo más conocido de aplicación del análisis factorial en psicología es el constituido por la teoría de las aptitudes primarias, conjunto de factores básicos de la capacitación intelectual. Además Thurstone publicó otros trabajos de interés psicométrico sobre calibración de items (1.947), métodos de puntuación de tests (1.919), fiabilidad y validez (1.931), la fórmula de Spearman-Brown (1.928), e incluso elaboró un manual de estadística (1.935).

4. Escalamiento Psicológico.

Las aportaciones de Thurstone son relevantes para la psicofísica como tal, y, a la vez, constituyen los primeros procedimientos bien definidos de escalamiento psicológico, de modo que una parte de la obra de Thurstone puede considerarse

propriadamente escalamiento psicológico. El "*método de las comparaciones apareadas*" estaba directamente importado y adaptado del campo psicofísico, pero es sin duda uno de los modelos tradicionales de escalamiento psicológico más valiosos. Entre otros métodos, Thurstone dio forma al método de escalamiento psicológico que se conoce como "*de los intervalos aparentemente iguales*". Este procedimiento permite elaborar escalas de opiniones y actitudes a través de una serie de fases que incluyen someter los primeros ítems seleccionados a un conjunto de jueces, para calibrar su posición en el continuo de actitud, antes de que puedan ser usados con los sujetos a medir. El procedimiento permite también efectuar una calibración precisa de los ítems y ofrece criterios para seleccionar razonablemente los elementos que han de formar la escala final.

En el método de los intervalos aparentemente iguales se escriben los estímulos, -ítems o enunciados sobre el objeto de actitud- a escalar, en fichas separadas. A cada sujeto que actúa como juez se le solicita que coloque cada ficha en uno de once montones, desde "ítem muy desfavorable" hasta "ítem muy favorable al objeto de actitud". Es decir, se le pide que atribuya a cada ítem una puntuación de 1 a 11 según el considere que refleja una actitud más o menos favorable. La media o la mediana de los puntos obtenidos por cada ítem en la muestra de jueces constituye el valor de escala del ítem. Después esos ítems, formando un cuestionario, pueden utilizarse para medir sujetos. En síntesis, si un sujeto J dice que está de acuerdo con un ítem A, y ese ítem A ha sido clasificado por los jueces, -en promedio-, como muy favorable al objeto de actitud, entonces podemos decir que el sujeto J presenta una actitud muy favorable.

Sin embargo, el método de los intervalos aparentemente iguales presentaba también algunas dificultades. Entre ellas, la necesidad de un proceso costoso de elaboración, el difícil supuesto de que la posición de los propios jueces en el continuo de actitud no influye en su contribución a la calibración, y la débil consideración de la cuestión de la dimensionalidad.

Remsis A. *Likert* en 1.932 facilitaría un procedimiento de escalamiento de actitudes que venía a solventar aceptablemente y con una extraordinaria sencillez las dos primeras dificultades mencionadas para el método de Thurstone. En el método de la *estrategia de "respuesta-directa" de Likert* los sujetos evalúan directamente sobre una escala de "anclajes" de acuerdo o desacuerdo su relación con un ítem. Por ejemplo, el sujeto contesta si esta "nada", "poco", "bastante" o "muy de acuerdo" con un enunciado referido al objeto de actitud.

Quizás se trate del modo más utilizado para la medida de las actitudes. El mismo Likert introdujo su procedimiento como respuesta a las complejidades, procesos laboriosos y numerosas asunciones que implicaban las técnicas de toma de datos para el escalamiento propuestas por Thurstone durante la década de los veinte. Likert probó empíricamente que su técnica arrojaba un cuestionario tan fiable como los elaborados por la técnica de Thurstone y que existía una alta correlación (en torno a 0'80) entre ambos procedimientos.

Sin embargo, Likert reconoció explícitamente las aportaciones de Thurstone. "La aplicación de sus escalas de actitudes [las de Thurstone] implicaba varios supuestos estadísticos. por ejemplo, que los valores de escala de los enunciados son independientes de la distribución de actitudes de los jueces que los evalúan-, supuestos que como señala Thurstone, no han sido verificados. Además, el método es excesivamente trabajoso. Parece legítimo preguntarse si en realidad funciona mejor que las escalas más sencillas que se pueden emplear y, al mismo tiempo, si no es posible construir escalas igualmente fiables sin hacer innecesarios supuestos estadísticos. [...] En el temor de que haya quienes interpreten erróneamente este artículo como un 'ataque' a los métodos de Thurstone, deseo subrayar enérgicamente que me limito a intentar llamar la atención sobre ciertos problemas metodológicos y que de ningún modo estoy convencido de que con los datos que presento se cierre la cuestión." (Likert, 1.932).

En realidad el trabajo de Likert venía a sumarse a una intensa actividad en torno a la cuestión de la medición de las actitudes y la personalidad que una serie de autores estaban desarrollando. En esta línea iban los trabajos de Bain, de Vetter, de Katz y Allport, y sobre todo una serie de trabajos de Murphy. La línea de trabajo que dio lugar al artículo de Remsis Likert había sido comenzada por el mismo Likert bajo la dirección de Gardner Murphy en 1.929. Los cuestionarios referidos a la medición de actitudes que presenta Likert en su artículo habían sido desarrollados por ambos. A pesar de las modestas pretensiones y de las prevenciones que Likert expresó acerca de su trabajo, éste ha pasado a constituir un punto de partida de un procedimiento sencillo y muy ampliamente utilizado en el campo de la medición de las actitudes y también en el campo de la medición de la personalidad.

En 1.936 Guilford publica 'Métodos Psicométricos' efectuando una primera síntesis de una década de desarrollos en el campo. Casi veinte años después, en 1.954, la obra fue revisada

y actualizada incluyendo un número de avances acerca de los métodos de escalación, pero con una atención limitada a las derivaciones y a los métodos de escalamiento multidimensional. Guilford ponía de relieve reiteradamente en aquel trabajo que la mayoría de las personas que se habían ocupado del desarrollo de la teoría de los tests habían ignorado o desconocido el trabajo en el campo del escalamiento psicofísico. Se realizaron algunos intentos de acercamiento entre los dos campos, entre los que pueden citarse los de Mosier en 1.940 y 1.941. Sin embargo, Gulliksen en 1.950 todavía consideraba la integración entre ambos campos de trabajo, teoría de test y escalamiento, como una meta futura que debía lograrse.

Los trabajos de Thurstone y de Likert habían dejado pendientes importantes cuestiones. Entre ellas la de la dimensionalidad de la escala y la de la interpretación de la puntuación total. Aquellos procedimientos habían sido concebidos suponiendo que la escala era unidimensional y orientando la selección de items de forma que presentaran una relación consistente con la escala. Louis Guttman enfocaría estas cuestiones desde el punto de vista opuesto.

Desde los principios de los años cuarenta Guttman publica una serie de trabajos que dan lugar a lo que se ha dado en conocer como *escalamiento de Guttman*, análisis de escalograma o modelo determinista de Guttman. El escalamiento de Guttman provee de instrumentos de medida donde, al menos en el caso ideal, la interpretación de la puntuación total es inequívoca en términos de qué afirmaciones o items concretos describen al sujeto en una dimensión única y acumulativa. Por ejemplo, si un sujeto obtiene una puntuación de 3 en una escala Guttman de siete items, ello significa que ha estado de acuerdo con los tres primeros items y en desacuerdo con los cuatro restantes. Los escalogramas de Guttman están concebidos como una "escalera" en la que para contestar afirmativamente o acertar el ítem dos hay que haber contestado afirmativamente o haber acertado el ítem uno. Está es su característica distintiva.

Los métodos de Guttman han sido objeto de una dilatada investigación que ha dado lugar a una variedad de procedimientos alternativos y de conexiones e interpretaciones en términos de otros modelos. En conjunto la obra de Guttman constituye un conjunto variado e innovador de respuestas al desafío de entender los datos y estructurar la comprensión de los mismos.

Concepción de la medición de Stevens. Fue en septiembre de 1.941 cuando en el "International Congress for the

Unity of Science" S.S. Stevens dió a la luz su famosa interpretación de la medición y sus conocidos cuatro niveles. Para Stevens medir consiste en asignar números a objetos según reglas. Hay cuatro niveles de medida: el nominal, donde los números actúan como etiquetas; el ordinal, donde los números expresan orden pero no se conoce la distancia entre ellos, no hay una unidad de medida; el de intervalo, donde los números representan orden y además hay una unidad de medida constante, pero no cero absoluto real; y el de razón, donde hay orden, unidad de medida y un cero absoluto. Según el nivel de medida en que esté una variable podrán o calcularse unos u otros estadísticos.

La concepción de la medición de Stevens era lo suficientemente amplia como para abarcar la mera asignación de números sin carácter ordinal (nivel nominal), pero a la vez, lo suficientemente estructurada como para definir que estadísticos resultaban permisibles (y, por tanto, cuales no) a cada nivel de medición. En realidad, el modo de Stevens de entender la medición y sus niveles no se extendió en el campo de las ciencias sociales hasta su "Handbook" de 1.951, pero a partir de ese momento puede decirse que ha constituido una escuela principal inevitablemente citada en el modo de entender la medición psicológica. La cuestión ha sido vuelta a considerar en la década de los ochenta, suscitando una polémica acerca de la cuestión de las restricciones y adscripción de estadísticos en función de las escalas de medida.

La obra de Torgerson. En 1.950 un Comité de Métodos y Teoría de Escalamiento presidido por Harold Gulliksen y en el que se encontraban Paul Horst (hasta 1.954), John E. Karlin, Paul F. Lazarsfeld, Henry Margenau, John Volkmann y Frederick Mosteller fue seleccionado por el "Social Science Research Council" para revisar el estado del arte en escalamiento en el marco de las ciencias sociales. Como un producto del trabajo del comité, se encargó a Warren S. Torgerson la revisión de la literatura acerca del escalamiento. Como explica el mismo Gulliksen (1.958), Torgerson fue escogido para este trabajo por su intensa dedicación al campo. Una dedicación que ya puso de manifiesto en su tesis doctoral que constituye uno de los importantes trabajos pioneros del escalamiento multidimensional.

El Comité planificó el desarrollo de una monografía sobre escalamiento, a modo de un manual con revisión e integración de la literatura, sin excluir nuevas aportaciones originales, que debía estar lista en 1.954. Sin embargo, Torgerson fue reclamado al servicio de la "Navy" desde 1.952 hasta 1.954, lo que desarticuló el plan de trabajo original y retrasó la obtención de

un texto otros cuatro años. Finalmente los trabajos concluyeron en Julio de 1.958 con la publicación por Warren S. Torgerson, entonces ya miembro del staff del Laboratorio Lincoln, en el Instituto de Tecnología de Massachusetts, de la obra "Theory and methods of scaling", que constituye, todavía actualmente, una pieza clave sobre la que se articula el desarrollo de los métodos de escalamiento. Además de los miembros del Comité, incluido Gulliksen, otros como Messick, Tucker, Keats, el mismo F.M. Lord y también C.H. Coombs estuvieron al tanto del desarrollo de esta obra y contribuyeron a la revisión de los capítulos.

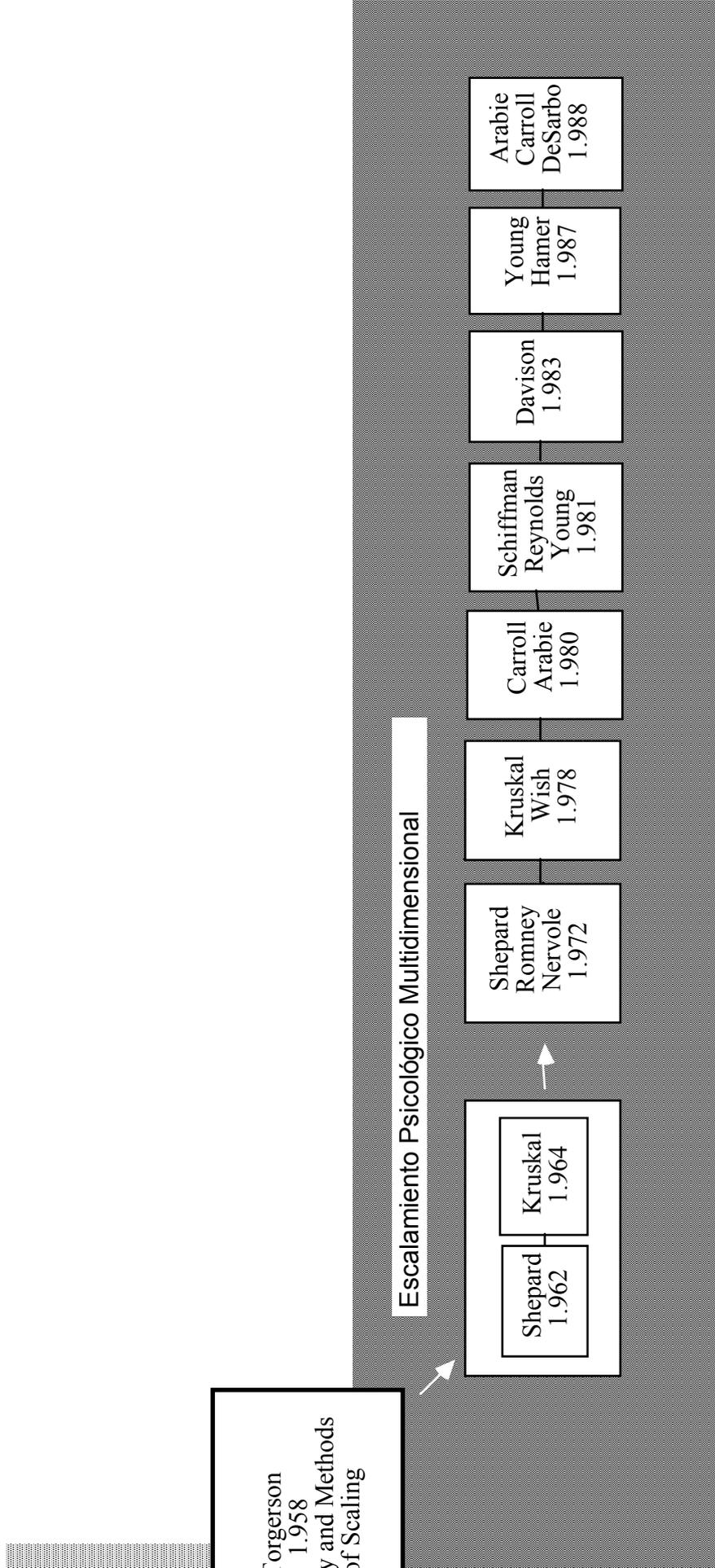
La obra *'Teoría y métodos de escalamiento'* de Torgerson, publicada en 1.958, dedica varios capítulos a la fundamentación de la medición y los métodos derivados del trabajo de Thurstone, pero también incluye capítulos acerca del escalamiento multidimensional, y los modelos probabilistas y deterministas para datos categoriales. El capítulo 3 de esta obra ofrece algunos ejes acerca de la clasificación de los métodos de escalación, como, por ejemplo, la distinción entre escalamiento de sujetos, de estímulos y la aproximación de respuesta, que siguen siendo utilizados por su capacidad para clarificar el desarrollo de la investigación en este campo (McIver y Carmines, 1.981).

Una de las más apreciables aportaciones conceptuales de la obra de Torgerson para el lector actual es que sugiere claramente la integración de los modelos deterministas, de los modelos de rasgo latente en el marco de los probabilísticos, y de los modelos de desplegamiento, en el conjunto de la teoría del escalamiento. De este modo la obra de Torgerson deja abiertos los puentes que enlazan el escalamiento con la teoría clásica de los tests y con la teoría de la respuesta al ítem. Desde este punto de vista, la obra de Torgerson es una contribución especialmente valiosa a la necesidad de integración que Guilford había puesto de manifiesto.

El trabajo de Torgerson excluye, sin embargo, un tratamiento explícito de técnicas multivariadas, como el análisis factorial, que, no obstante, como el mismo Torgerson señala, pueden usarse bajo conjuntos de supuestos delimitados como modelos de medida. En síntesis, la teoría y métodos de escalamiento de Torgerson puede entenderse que constituye el eslabón central entre los procedimientos unidimensionales de escalamiento, en su mayor parte desarrollados con anterioridad a la obra, y los multidimensionales, nacidos antes de esta obra central pero desarrollados muy especialmente después de la misma.

En la *Figura 11* se aprecia este reconocido papel central de la obra de Torgerson (1.958). En la parte superior de la misma se recogen algunas de las aportaciones y fechas significativas del escalamiento unidimensional; en la parte inferior se presentan algunas obras importantes -trabajos originales o de carácter docente- para la comprensión del escalamiento multidimensional. La *figura 11* sugiere como la obra de Torgerson, situada poco después del medio siglo, puede considerarse una pieza central de valor no sólo histórico.

UNAS OBRAS RELEVANTES DEL ESCALAMIENTO PSICOLOGICO.



Torgerson
1.958
Psychometric Methods
of Scaling

Shepard
1.962

Kruskal
1.964

Shepard
Romney
Nervele
1.972

Kruskal
Wish
1.978

Carroll
Arabie
1.980

Schiffman
Reynolds
Young
1.981

Davison
1.983

Young
Hamer
1.987

Arabie
Carroll
DeSarbo
1.988

Escalamiento Psicológico Multidimensional

El Diferencial Semántico de Osgood. Un año antes de la publicación de la obra de Torgerson, en 1.957, Osgood, Suci y Tannenbaum, dieron a la luz un nuevo procedimiento de medición, el *diferencial semántico*, que llegaría a alcanzar una gran divulgación en el campo de las ciencias sociales. En realidad el procedimiento puede considerarse una extensión, con algunas complicaciones adicionales, del escalamiento de Likert, fundado sobre la definición de los objetos de interés bajo un conjunto de continuos bipolares definidos por adjetivos opuestos. En la técnica del diferencial semántico se pide a cada sujeto que califique el estímulo o ítem de interés en una serie de adjetivos bipolares (frio-caliente, duro-blando, etc.). Cada par de adjetivos ocupan los extremos de una escala de siete puntos. Así por ejemplo, en un estudio sobre las actitudes ante la guerra se le puede pedir a un sujeto que califique el ítem "gasto militar" en cada uno de esos adjetivos. Cada ítem queda descrito en cada escala bipolar. Todos los ítems son descritos en el mismo conjunto de escalas de adjetivos bipolares. Después, usualmente, las puntuaciones en esos adjetivos son agrupadas de algún modo prefijado en factores o dimensiones. (Esas dimensiones suelen provenir de resultados de análisis factoriales previos.) La técnica del diferencial semántico todavía goza de un amplio uso en el campo de la medición de actitudes en psicología social.

Luce en 1.959 presentó un modelo general de conducta de elección individual. En 1.961 F. Restle publicó su "Psicología del Juicio y Elección" donde desarrolló un modelo formal, utilizando teoría de conjuntos y cálculo de probabilidades, de la probabilidad de la respuesta de un sujeto en situaciones donde un número de alternativas están presentes y delimitadas. Como principales supuestos estableció los de equiprobabilidad e independencia. Como elementos influyentes en la decisión consideró los aspectos valorables de las alternativas. En la medida en que los aspectos valorables de las alternativas fuesen distintos, de diferentes cualidades, se produciría impredecibilidad en la elección. Si las alternativas no poseen aspectos valorables comunes el modelo de Restle aparece consistente con el de Luce. Una de las razones que vuelven importantes los trabajos de Luce y Restle es su extensión al análisis de los juicios de similitudes. El mismo Thurstone (1.959) resumirá buena parte de lo esencial de su obra en "The measurement of values". También ese año Stevens añadió a sus cuatro niveles de escalas clásicos una quinta la de "intervalo logarítmico" que nunca llegaría a adoptarse dentro del conocimiento general y que pronto se identificó como asimilable a su anterior nivel de escala de intervalos.

A principios de la década de los sesenta Gulliksen y Messick publican "Psychological Scaling: Theory and Applications" (1.960) y C.H. Coombs (1.960) da a la luz una versión más acabada de su teoría de los datos. Esta teoría "de los datos" enlaza con la cuestión clásica de las escalas de medida, pero enfrenta el problema de un modo original. Desde su teoría, los modelos de medida pueden ser clasificados sugiriendo las técnicas de escalamiento oportunas.

Teoría del desplegamiento unidimensional de Coombs. Como un fruto importante de este trabajo de Coombs en el campo de la teoría de los datos, había aparecido durante la década anterior la *teoría del desplegamiento unidimensional*, anticipada en un trabajo (Coombs, 1.950) titulado "Escalamiento psicológico sin una unidad de medida". El modelo de desplegamiento de Coombs nunca ha llegado a hacer excesiva fortuna en el trabajo empírico de los científicos de la conducta. Sin embargo, es un enfoque poderosamente atractivo y original de la cuestión cuyos frutos aplicados quizás estén aún por darse. El modelo de desplegamiento unidimensional de Coombs permite entender como una misma escala común (que denomina escala J, junta o común) puede ser vista de forma diferente por diferentes sujetos (escalas I, individuales o particulares de los sujetos) en función de su perspectiva. Una misma escala o continuo psicológico da lugar a diferentes respuestas por diferentes sujetos en función de su perspectiva individual, en función de su posición en ese continuo.

Coombs presenta una visión más completa de su método de escalamiento en 1.964, entre tanto Bennett y Hays (1.960) extendieron la técnica de escalamiento unidimensional de Coombs formalizándola para el caso multidimensional, en tanto que Coombs y Kao (1.955) habían sentado las bases para la conexión entre el análisis factorial y el desplegamiento multidimensional.

Contribuyendo al desarrollo de los métodos de escalamiento de Guttman, Willis (1.961) sugiere incluir dos cuestiones para cada ítem, una relativa al acuerdo o desacuerdo, la otra a la intensidad asociada. Utilizando esta información Willis consigue una alta reproductibilidad. En esta línea, contemporáneamente, Heath (1.961), tomando en cuenta la magnitud de las discrepancias con la escala, desarrolla dos índices de reproductibilidad.

En 1.963 Ledyard R. Tucker realiza una revisión del campo del escalamiento y de la teoría de los tests distinguiendo tres grandes áreas: primero, el escalamiento psicológico, segundo

la teoría de los tests, y, en tercer lugar, el análisis factorial. En 1.965 Ekman y Sjöberj publican la primera revisión del *Annual Review* especialmente dedicada a escalamiento. Antes de esa fecha la revisión más extensiva anterior debe buscarse en el trabajo de Schiffman y Messick de 1.963. En el 1.969 la revisión de Zinnes tomará la cuestión del juicio como eje conceptual ordenador.

Desde principios de la década de los sesenta el trabajo en escalamiento se orienta cada vez más hacia el *escalamiento multidimensional (MDS)*, sin olvidar algunas aportaciones a los métodos de escalamiento unidimensional. Los manuales y las revisiones estarán a partir de este punto más orientadas hacia el MDS. Sin embargo, existen algunas excepciones importantes. Puede mencionarse entre ellas la obra de Bock y Jones "La medición y predicción del juicio y la elección" publicada en 1.968, que se concentra en el análisis de los modelos de Thurstone y derivados sin considerar el análisis multidimensional. Esta obra ha sido considerada como una continuación, en lo relativo a los modelos derivados de Thurstone, de la importante obra de Torgerson publicada una década antes. También ha tenido una especial divulgación "Fundamentals of Scaling and Psychophysics" de Baird y Noma, publicada en 1.978. Ese mismo año Shye se ocupará de presentar una visión integrada de los desarrollos y complicaciones posteriores de los métodos de Guttman en su "Theory construction and data analysis"

5. Escalamiento Multidimensional.

Origen. El origen del escalamiento multidimensional puede ubicarse en los trabajos de Young y Householder y de Richardson, de 1.938. Sin embargo, suelen tomarse los trabajos de Shepard de 1.962 y de Kruskal de 1.964 como el origen reciente de los amplios desarrollos y abundante literatura que sobre este conjunto de técnicas se ha publicado en estas últimas décadas. Si bien Torgerson ya se había ocupado del tema en 1.958, es a partir de los trabajos de Shepard y Kruskal cuando se origina una extensa investigación que ha dado lugar a un buen número de obras importantes.

Concepto. Los procedimientos de escalamiento multidimensional pueden considerarse extensiones más o menos sofisticadas de las técnicas de escalamiento unidimensional a situaciones en las que los estímulos pueden ser descritos más adecuadamente en un espacio k-dimensional. Según Torgerson (1.958) *el problema típico que debe enfrentar el escalamiento*

multidimensional puede describirse así: "dado un conjunto de estímulos que varían con respecto a un número desconocido de dimensiones, determinar (a) la dimensionalidad mínima del conjunto, y (b) las proyecciones del estímulo (valores de escala) en cada una de las dimensiones implicadas."

Esto implica describir a cada objeto de estudio mediante un vector de dimensión k , de modo que para describir un objeto dado son necesarios tantos valores, proyección sobre los ejes dimensionales, como dimensiones presenta el modelo. La distancia entre dos puntos en el espacio es función del grado de similitud entre los objetos considerados.

Distancias. El escalamiento multidimensional implica un modelo geométrico espacial y un modelo de distancias, una teoría que vincule las distancias entre puntos con relaciones observables, que establezca cómo son obtenidos los datos de distancias y sus significados.

Desde el punto de vista del escalamiento multidimensional, dado un conjunto de distancias o disimilaridades entre puntos -objetos, estímulos, etc.- puede considerarse si es posible representarlas adecuadamente bajo alguno de estos modelos, y después, determinar la dimensionalidad del espacio. Torgerson ha resumido e integrado un conjunto de teoremas de Young y Householder (1.938) que permiten enfrentar estas cuestiones. Por su parte, los modelos de distancias establecen como se determinan las distancias con que operan los modelos espaciales. Usualmente una distancia es una representación o medición cuantitativa de la disimilaridad o alejamiento entre dos "objetos" comparados en alguna característica. Por ejemplo, el coeficiente de correlación de Pearson entre dos variables puede ser considerado como una medida de similitud o proximidad entre ambas; el mismo coeficiente descontado de la unidad se convierte en una medida de disimilaridad o alejamiento. Torgerson (1.958) pone de manifiesto que los modelos de distancia son análogos o pueden obtenerse como extensión de los modelos de escalamiento unidimensional.

El modelo espacial más utilizado ha sido el euclidiano, invariante frente a transformaciones euclidianas que implican traslación o rotación ortogonal de los ejes, donde la distancia entre dos puntos se representa por:

$$d_{jk} = \sqrt{\sum_{m=1}^r (a_{jm} - a_{km})^2} \quad (j, k = 1, 2, \dots, n)$$

aunque también se han postulado espacios no euclídeos, como el "city-block" o distancia Manhattan propuesto por Attneave (1.950) donde:

$$d_{jk} = \sum_{m=1}^r |a_{jm} - a_{km}| \quad (j, k = 1, 2, \dots, n)$$

en el que las distancias no permanecen invariantes frente a la rotación de los ejes. La distancia Manhattan y la distancia euclídea son dos casos particulares de la clase de distancias o disimilaridades conocidas como métricas de Minkowski que responden a la forma general:

$$d_{jk} = \sqrt[n]{\sum_{m=1}^r (|a_{jm} - a_{km}|)^n} \quad (j, k = 1, 2, \dots, n)$$

De modo que si $n=1$ esta ecuación equivale a la de Attneave, y si $n=2$ equivale al modelo euclídeo.

Estas medidas de distancia o disimilaridad, junto con otras como la D^2 de Mahalanobis (1.936), que no son métricas de Minkowski, constituyen uno de los grupos de medidas de similaridad utilizadas en el *análisis cluster*. El análisis cluster o análisis de conglomerados es un conjunto de técnicas estadísticas que permiten agrupar o clasificar objetos, sean sujetos o variables, en función de su semejanza. Esta conexión entre las técnicas cluster y el escalamiento multidimensional no resulta sorprendente si se admite la definición que hace Everitt (1.980) del concepto de clusters como "regiones continuas de un espacio que contiene [regiones con una] densidad relativamente alta de puntos, separadas de otras regiones semejantes por regiones que contienen una densidad relativamente baja de puntos". La función del análisis cluster consistiría precisamente en determinar esas regiones y clasificar a los objetos bajo estudio en función de las mismas.

Principalmente en la década de los setenta y en la de los ochenta, acorde con un incremento sustancial en los artículos de investigación y aplicación acerca del MDS, se han desarrollado un buen número de exposiciones sistemáticas y manuales del MDS, con orientaciones más o menos teóricas o aplicadas. Quizás deba destacarse primero la obra de dos volúmenes "Multidimensional Scaling: Theory and applications in the behavioral sciences", editada en 1.972 por Shepard, Rimney y Nerlone. En 1.978 Kruskal y Wish publicaron su muy breve y didáctico "Multidimensional Scaling", que puede considerarse una

introducción docente al campo. El puesto de manual docente lo disputa también "Introduction to Scaling" de Van der Ven, publicado en 1.980. De mayor entidad teórica es la obra de Schiffman, Reynolds y Young de 1.981, que presenta la ventaja de ser extraordinariamente útil al investigador que ha de utilizar MDS mediante alguno de los principales programas disponibles. También merece una atención destacada la revisión y ordenación del campo que aparece en el manual de Davison (1.983) y el libro básico de Dunn-Rankin (1.983), que trata también el escalamiento unidimensional y el análisis cluster. Por último mencionaremos también la obra de Arabie, Carroll y DeSarbo publicada en 1.988 y que puede considerarse una extensión avanzada del trabajo de Kruskal y Wish. Una aproximación al desarrollo histórico y principales elementos del MDS y su aplicación puede encontrarse en los trabajos de Coxon (1.982), así como en los mencionados de Kruskal y Wish (1.978) y de Schiffman, Reynolds y Young (1.981), una más sucinta presentación introductoria puede encontrarse en el mencionado artículo de Davison (1.983).

En 1.983 la revista "Applied Psychological Measurement" dedicó un número monográfico al MDS titulado "Multidimensional Scaling and its Applications", editado por Mark L. Davison y Lawrence E. Jones. El primer editor realizó una introducción al número en la que prestaba especial atención al desarrollo del campo durante las dos décadas anteriores, con el valor de revisión histórica que ello tiene. En el mismo volumen L.G. Cooper, L.E. Jones, E.J. Shoben, y J.B. Rounds con M.A. Zevon se ocuparon respectivamente de revisar los desarrollos del MDS en los campos del marketing, la percepción social, cognición y conducta, la psicología cognitiva y la psicología vocacional. Los trabajos de W.J. Heiser con J. Meulman y de I. Spense se ocuparon respectivamente del escalamiento multidimensional con constricciones, incluida la confirmación, y de estudios de simulación. En conjunto estos trabajos dan muestra de la vitalidad de la investigación en el desarrollo del MDS, del interés que presenta para el campo de la medición en ciencias comportamentales y de la extensa y variada gama de aplicaciones que está alcanzando.

Definición de MDS. Desde la primera revisión de Ekman y Sjöberg sobre el campo del escalamiento, el MDS había aparecido como un apartado dentro del conjunto en las revisiones sucesivas sobre escalamiento. Así sucede por ejemplo en Zinnes (1.969) y en Cliff (1.973). Carroll y Arabie (1.980) presentan la primera revisión del "Annual" dedicada exclusivamente al escalamiento multidimensional. Este trabajo ha tenido un impacto especial al contribuir a definir el MDS y al ensayar una taxonomía

del campo. Carroll y Arabie distinguen dos aproximaciones a la definición del MDS. Según una de ellas el MDS englobaría todas las técnicas para estimar parámetros en modelos geométricos para ofrecer una representación de la estructura de los datos. Bajo esta definición, que no es la más utilizada, el MDS incluiría el análisis cluster, el análisis factorial y el análisis discriminante. Con un significado más restringido, el MDS puede entenderse referido al *conjunto de técnicas estadísticas para estimar los parámetros que evalúan el ajuste de varios modelos de distancias espaciales para datos de proximidad o preferencia, y la representación mediante coordenadas de la estructura de estímulos que resulta de la aplicación de esas técnicas* (Davison, 1.983). Posteriormente, otras revisiones como la Young de 1.984 o las contenidas en los principales manuales antes citados han recogido sucesivas aportaciones al campo.

En síntesis, el escalamiento psicológico, nacido de los métodos de escalamiento psicofísicos, se ha desarrollado desde métodos unidimensionales que al principio no consideraban en absoluto la cuestión de la dimensionalidad, hacia métodos multidimensionales que permiten representar métricamente en un conjunto de dimensiones. En su día los métodos de escalamiento multidimensionales fueron la lógica extensión de los métodos unidimensionales, actualmente más bien los métodos unidimensionales pueden ser vistos como un caso particular del escalamiento multidimensional. En conjunto, los procedimientos de escalamiento constituyen uno de los grandes brazos históricos de la medición psicológica, como ya pusiera de manifiesto Torgerson. Un brazo que, según el proyecto de Gulliksen, debía llegar a confluir y unificarse con los desarrollos de la teoría de los tests mentales.