

Métodos de escalamiento aplicados a la priorización de necesidades de formación en organizaciones

Susana Sanduvete Chaves, M^a Isabel Barbero García*, Salvador Chacón Moscoso, J. Antonio Pérez-Gil, F. Pablo Holgado Tello*, Milagrosa Sánchez Martín y J. Antonio Lozano Lozano
Universidad de Sevilla y * Universidad Nacional de Educación a Distancia

Los criterios para priorizar las necesidades que justifican las acciones formativas a implantar no se suelen explicitar a priori en los programas de formación continua en contextos organizacionales. En este trabajo se proponen los métodos de escalamiento como procedimiento factible y útil para identificar criterios explícitos de priorización de necesidades, y se concreta cuál de ellos es más apropiado en este contexto de intervención. 404 empleados de una organización pública cumplimentaron un cuestionario ad hoc para priorizar necesidades formativas en diferentes áreas durante el período 2004 al 2006; concretamente, se ordenaron 117, 75 y 286 estímulos, respectivamente. Se calcularon y compararon las ordenaciones obtenidas con cuatro métodos de escalamiento: el método de Dunn-Rankin y tres métodos derivados de la Ley del Juicio Categórico de Thurstone, concretamente ordenación por rangos, intervalos sucesivos e intervalos aparentemente iguales. Con los resultados, se constata la factibilidad y utilidad de estos métodos de escalamiento para la solución de los problemas planteados; a partir de los métodos más precisos, se propone el uso del método de ordenación por rangos por su parsimonia (respecto a sencillez en su procedimiento); y se anticipan futuras líneas de actuación.

Scaling methods applied to set priorities in training programs in organizations. Criteria to assess the needs in order to plan training programs are not usually defined explicitly in organizational contexts. We propose scaling methods as a feasible and useful procedure to set priorities of training programs; also, we propose the most suitable method for this intervention context. 404 employees from a public organization completed an ad hoc questionnaire to assess training needs in different areas from 2004 to 2006; concretely, 117, 75 and 286 stimuli were scaled, respectively. Then, four scaling methods were compared: Dunn-Rankin's method and three methods derived from Thurstone's Law of Categorical Judgment —ranking, successive intervals and equal-appearing intervals. The feasibility and utility of these scaling methods to solve the problems described is shown. Taking into account the most accurate compared methods, we propose ranking as the most parsimonious method (with regard to procedure simplicity). Future research developments are described.

La evaluación de necesidades puede ser definida como un procedimiento sistemático por el que se identifican y priorizan las carencias de un colectivo y se planifican las líneas de acción para eliminar o reducir dichas carencias (Witkin y Altschuld, 1995). Se considera una fase crítica previa a toda intervención, puesto que (Anguera, Chacón y Blanco, 2008): (a) determina la pertinencia de las necesidades planteadas; (b) delimita los criterios para priorizar necesidades; (c) aporta mediciones de las necesidades delimitadas antes de la intervención; (d) justifica la asignación de los recursos (usualmente limitados) a las diferentes acciones; y (e) aporta criterios de inclusión/exclusión de los participantes a las condiciones del programa.

En el ámbito de la formación continua, considerada hoy día uno de los pilares en los que se sustentan los Sistemas de Gestión de la Calidad en las organizaciones (Gómez y Palací, 2003), las tomas de decisión final respecto a los planes de formación a desarrollar se basan, principalmente, en los criterios establecidos por sus responsables: políticos en las organizaciones públicas y propietarios/gerentes en las organizaciones privadas. Como consecuencia, estos criterios de valoración no se delimitan explícitamente antes de la recogida de datos y suelen ser divergentes en función del grupo de implicados que los establezca, por disparidad de intereses (Chacón, Anguera, Pérez-Gil y Holgado, 2002).

En una revisión de la literatura, hasta junio de 2008, se constató que la evaluación de las necesidades formativas suele presentar las siguientes debilidades (Sanduvete, 2008): (a) sólo el 6% de los estudios definían los constructos de manera conceptual (debilidad relacionada con la pertinencia) y empírica (en relación con los criterios de priorización de necesidades); (b) sólo en el 32% de los casos se realizaba al menos una medida antes y otra después de la intervención/formación (debilidad relacionada con la identificación de necesidades previas a la intervención); (c) en el 50% de los

casos no se explicitaban los recursos asignados en el proceso de formación (debilidad ligada a la justificación de la asignación de recursos a diferentes acciones formativas); y (d) en el 70% de los casos no se explicitaban los criterios de selección de la muestra (debilidad en relación con la inclusión/exclusión).

A pesar de la conveniencia de presentar posibles soluciones a los problemas del proceso evaluativo descritos, la mayor parte de las evaluaciones de necesidades en este contexto se basan en datos meramente descriptivos (Chacón, Lara y Pérez-Gil, 2002; Wen y Gustafson, 2004). Mientras, en otros contextos de intervención como la actuación ante situaciones de emergencia (White, Turoff y Van de Walle, 2007) o en el ámbito de la salud (Wen y Gustafson, 2004), los métodos de escalamiento han sido aplicados en priorización de necesidades con resultados satisfactorios.

Los objetivos generales de este trabajo son: (a) constatar la factibilidad y utilidad de los métodos de escalamiento como procedimiento adecuado para obtener criterios explícitos de priorización de necesidades en el ámbito de la formación en las organizaciones; y (b) concretar el método más adecuado en nuestro ámbito: partiendo de niveles de precisión similar entre métodos, seleccionar el más parsimonioso en cuanto a sencillez en su procedimiento.

En este sentido, la factibilidad es entendida como la posibilidad de implementar los métodos de escalamiento en el contexto de la formación de las organizaciones (Anguera et al., 2008); la utilidad, como la influencia de los resultados de los métodos de escalamiento (priorización de necesidades formativas) en la posterior toma de decisión sobre el programa a implementar (Vedung, 1995).

Por otra parte, en la literatura no se plantean diferencias importantes entre los métodos de escalamiento respecto a su fiabilidad o validez (Meliá, Sospedra, Molina, Ramón e Islas, 1991; Meliá, Sospedra, Ramón y Molina, 1992); así que, una vez seleccionados los métodos de escalamiento que podamos aplicar en el contexto de este trabajo, se elegirá, de entre aquellos con similar nivel de precisión, el que presente el procedimiento de aplicación más sencillo, con el fin de tratar de agilizar su uso y tener menos probabilidad de cometer errores en su aplicación.

El ámbito de la formación en las organizaciones que contextualiza este trabajo implica, básicamente, las siguientes condiciones: (a) la ordenación de un amplio número de estímulos: se solicitan muchas acciones diferentes porque la formación es impartida en un amplio, variado y extenso ámbito geográfico, a participantes que presentan gran diversidad funcional en sus puestos de trabajo; y (b) una priorización de necesidades formativas como resultado: la ordenación precisa, tras la valoración comparativa por parte de los participantes, de un amplio conjunto de estímulos; concretamente, se ordenarían todas las acciones solicitadas, para posteriormente poder incluir en el plan de formación aproximadamente (en función de los recursos disponibles), las 100 valoradas como más relevantes.

Concretamente, se comparan cuatro métodos de escalamiento a partir de datos de necesidades de formación recogidos durante tres años en un organismo público: (a) el método de Dunn-Rankin (Arce, 1994) y tres derivados de la Ley del Juicio Categórico de Thurstone (Barbero, 1993); (b) el método de ordenación por rangos; (c) el método de intervalos sucesivos; y (d) el método de intervalos aparentemente iguales.

La elección de estos métodos y no otros se debió a su adecuación a las condiciones del contexto previamente mencionadas: (a) son escalas de juicio; (b) parten de un modelo de escalamiento centrado en los estímulos, lo que implica que las variaciones en-

contradas en las reacciones de los participantes ante los cursos de formación (estímulos) se deben a las diferencias de éstos respecto al grado de importancia para cubrir la necesidad (atributo psicológico); (c) presentan relación de dominancia entre estímulos (Coombs, 1964); y (d) permiten el escalamiento de un amplio número de estímulos sin conllevar por ello complicaciones excesivas en su procedimiento (Arce y Andrade, 2000; Meliá et al., 1992).

Por último, de acuerdo con los hallazgos de estudios previos sobre estos métodos, se espera encontrar: (a) correlación perfecta entre las ordenaciones obtenidas con los métodos de intervalos sucesivos y aparentemente iguales (Ponsoda, 1986); (b) alta correlación entre los resultados obtenidos con los métodos de Dunn-Rankin e intervalos aparentemente iguales (Meliá et al., 1991, 1992); y (c) mayor complejidad en el procedimiento del método de los intervalos sucesivos (Cuesta, Jiménez, Álvarez, Díaz y Paz, 1997).

Método

Participantes

Conforman la muestra 404 empleados de la Diputación de Sevilla, potenciales participantes en su Plan Agrupado de Formación Continua, que priorizaron sus necesidades de formación a lo largo de los años 2004 a 2006. La recogida de información se realizó de manera anónima por imperativo de la organización al considerarse que, de este modo, la participación aumentaría y disminuiría la deseabilidad social en la respuesta. Concretamente, la muestra fue conformada por 9 personas en el año 2004, 21 en el 2005 y 374 en el 2006. Se atribuye el incremento de participantes principalmente a que en el primer año sólo respondieron los responsables de formación de los municipios adheridos al plan como representantes de los demás trabajadores mientras que, en años sucesivos, pudieron responder directamente todos los potenciales participantes. Otras razones del aumento de participación son la mejora de las vías de comunicación y la consolidación del nuevo procedimiento para priorizar necesidades hasta ser percibido como un proceso natural dentro de la organización.

Instrumentos

Cada año, se elaboró un cuestionario de priorización de necesidades diferente, consistente en un listado de acciones formativas agrupadas, en función de la temática tratada, en once áreas de conocimiento publicadas anualmente en el Boletín Oficial del Estado por el Instituto Nacional de Administraciones Públicas. El número de acciones formativas priorizadas y su distribución por áreas de conocimiento en los distintos años se presentan en la tabla 1. La casuística de las acciones incluidas en los distintos años es variada; en todo caso, el criterio de inclusión se basaba en las demandas de los participantes. El listado completo de acciones formativas puede consultarse en <http://innoevalua.us.es/wakka.php?wakka=ArticulosPublicados&v=1eek>.

Las instrucciones escritas, comunes en los tres cuestionarios elaborados, fueron: «A continuación, aparece el listado de las posibles acciones formativas, divididas por áreas de conocimiento, que se han solicitado para el nuevo plan de formación. Para determinar la importancia que para usted tiene cada acción formativa, le rogamos que las ordene dentro de su área de conocimiento correspondiente (1º para la más importante, 2º para la siguiente, y así sucesivamente), sin repetir numeración». Adjunto al cuestionario se

presentó una ficha técnica de cada acción formativa donde se explicitó, al menos, las características de los destinatarios y la utilidad para el puesto de trabajo (Chacón, López y Sanduvete, 2004).

Procedimiento

De acuerdo al modelo teórico utilizado (Chacón, Lara y Pérez-Gil, 2002; adaptado de Witkin y Altschuld, 1995), se elaboró cada año un cuestionario de priorización de necesidades final en tres fases: (a) se recogió la información disponible de años anteriores acerca de las necesidades de formación de los participantes en planes previos; (b) con la información anterior, se elaboró un cuestionario que incluía un listado de acciones formativas y se solicitó a los participantes que marcaran las acciones necesarias para mejorar su rendimiento y que propusieran nuevas acciones; y (c) se obtuvo el cuestionario final consistente en un listado en el que se recogieron todas las acciones formativas detectadas en la fase anterior, asignadas a las distintas áreas de conocimiento por parte de los mismos participantes, por su contenido (aunque posteriormente fue revisado por expertos de cada área de conocimiento). Sólo quienes participaron en la fase anterior tuvieron la posibilidad de responder al cuestionario final, de manera anónima, ordenando las acciones por cada área en función de su importancia.

Los cuestionarios se distribuyeron entre los responsables de formación en el año 2004 y entre los potenciales participantes en el plan de formación en los años sucesivos, a través de dos vías: en papel por correo ordinario dirigido a su lugar de trabajo; y en la página web de la Sección de Formación Continua de la Diputación de Sevilla (<http://www.dipusevilla.es/formacioncontinua>). Tras la fusión en un único archivo de los datos recogidos por ambas vías, los valores perdidos fueron sustituidos por el último puesto posible (Little y Rubin, 2002).

Análisis de datos

Al tratarse de un procedimiento novedoso que pretendía ser desarrollado de manera independiente por personal interno de la propia organización, era necesario tener certeza de que la interpretación del procedimiento de los diferentes métodos, la elaboración de la hoja de cálculo y el registro de datos se realizaban adecuadamente. Por este motivo, dos codificadores independientes crearon sendas hojas de cálculo, registraron los datos y obtuvieron los valores escalares y los órdenes aplicando los cuatro métodos de escalamiento. Posteriormente, para estudiar el grado de acuerdo entre codificadores, se obtuvo la fiabilidad entre las ordenaciones resultantes con el coeficiente de correlación de rango o ρ de Spearman (Siegel, 1991).

Para determinar si existieron diferencias entre los resultados obtenidos con los diferentes métodos, se estudiaron las correlaciones de rango bivariadas de los valores escalares por un lado y de las ordenaciones por otro, de manera global con todos los estímulos, y de forma ilustrativa en varias áreas concretas.

Por último, con el fin de determinar la precisión de los diferentes métodos, se contabilizó la frecuencia con la que dos estímulos recibían la misma ordenación en cada método y se llevó a cabo la *prueba de rangos señalados y pares igualados de Wilcoxon* (Siegel, 1991), para concluir si las diferencias entre métodos en relación al número de empates obtenidos fueron significativas.

Resultados

La fiabilidad de la ordenación obtenida por los dos codificadores independientes obtuvo un coeficiente ρ de Spearman de 0.96, resultado altamente satisfactorio (las pequeñas diferencias fueron debidas a errores tipográficos por omisión de valores).

Tras el cálculo de los valores escalares y la ordenación de cada uno de los estímulos en las diferentes áreas por el método de Dunn-Rankin y los métodos de ordenación por rangos, intervalos aparentemente iguales e intervalos sucesivos de Thurstone se constató que, con este último método mencionado, los valores escalares no pudieron ser calculados para los estímulos posicionados en las categorías extremas al faltar datos necesarios para ello, puesto que el primer y el último intervalo del continuo quedaban abiertos: no existía límite inferior ni superior respectivamente y, en consecuencia, ninguno de los dos intervalos presentaba un valor de amplitud concreto.

Para no perder información, se optó por calcular en estos casos un valor escalar «hipotético», que daba la posibilidad de ordenar los estímulos que se encontraban en las categorías extremas; concretamente, para la categoría más baja se tomó como límite inferior un valor menor que el límite inferior de la siguiente categoría y para la categoría más elevada se tomó, como límite superior, un valor más elevado que el último límite conocido; una vez delimitados los extremos de las categorías, fue posible el cálculo de la amplitud estimada y, finalmente, del valor escalar.

Para estudiar las coincidencias entre los resultados encontrados en los diferentes métodos, en la tabla 2 se muestra la matriz de *correlaciones bivariadas*. Los valores presentados en la parte superior de la diagonal principal hacen referencia a las correlaciones obtenidas a partir de los valores escalares, mientras que las de la parte inferior de la diagonal son las correlaciones obtenidas a partir de las ordenaciones.

Todas las correlaciones son estadísticamente significativas tanto a nivel de confianza del 95% como del 99%, siendo más altas las obtenidas a partir de las ordenaciones. Los valores escalares

Tabla 1
Acciones formativas priorizadas por áreas de conocimiento en los distintos años

AÑO	A	B	C	D	E	F	G	H	I	J	K	TOTAL
2004	4	4	65	0	2	2	25	2	4	0	9	117
2005	2	5	37	0	2	5	15	0	3	2	4	75
2006	12	19	136	4	4	33	41	7	10	4	16	286

Áreas formativas: A= Dirección y Gerencia Pública; B= Económico Presupuestaria; C= Específico de Determinados Colectivos; D= Idiomas/Lengua; E= Información y Atención al Ciudadano; F= Jurídico - Procedimental; G= Nuevas Tecnologías. Información y Comunicación; H= Prevención de Riesgos Laborales; I= Recursos Humanos; J= Unión Europea; y K= Urbanismo y Medio Ambiente

obtenidos con el método de intervalos aparentemente iguales son los que más correlacionan con los obtenidos con los demás métodos, especialmente con el método de Dunn-Rankin. Las ordenaciones obtenidas con el método de Dunn-Rankin y ordenación por rangos presentan una correlación perfecta, y lo mismo ocurre con las obtenidas por el método de intervalos sucesivos y de intervalos aparentemente iguales.

La figura 1 presenta las diferencias en las ordenaciones en función de los métodos utilizados, incluyendo los 478 estímulos escalados a lo largo de los años. El eje de abscisas representa el orden obtenido por cada estímulo con los métodos de intervalos sucesivos (IS) e intervalos aparentemente iguales (IAI), puesto que eran coincidentes; y el eje de ordenadas, el orden coincidente obtenido en los métodos de Dunn-Rankin (DR) y ordenación por rangos (OR). Un punto en la diagonal trazada implica, por lo tanto, la coincidencia en los órdenes asignados a partir de los cuatro métodos; los puntos más alejados serían reflejo de los estímulos cuya ordenación difiere más entre pares de métodos.

De manera genérica, a lo largo de los años la tendencia en los resultados se mantuvo similar. La razón por la que existe una acumulación mayor de acciones formativas en las primeras posiciones es que no todas las áreas contenían un número elevado de acciones a escalar, por lo que los primeros puestos fueron cubiertos en todas las áreas, pero no los últimos.

Los resultados muestran un ajuste muy bueno entre las ordenaciones obtenidas con los diferentes métodos ($\rho^2= 0.966$). Para de-

terminar si las discrepancias entre métodos se daban homogéneamente a lo largo de los diferentes órdenes, se dividió dicho rango en tres partes iguales y se hallaron, en cada una de ellas, las correlaciones al cuadrado, constatándose así que las mayores diferencias se dan en las posiciones intermedias ($\rho^2= 0.496$).

Haciendo un estudio pormenorizado por áreas de conocimiento y diferentes años, los resultados se muestran usualmente en la misma línea; a modo ilustrativo, en las figuras 2, 3 y 4 se presentan las diferencias en las ordenaciones en función de los métodos utilizados, en el año 2006, en las áreas formativas C, F y G, respectivamente. Todas presentan altas correlaciones globales. Para marcar las diferencias en los extremos inferiores y superiores se consideraron como límites los percentiles 27 y 73 (Barbero, 1993). Las correlaciones específicas más altas se muestran en los primeros órdenes y las más bajas en las posiciones intermedias, aunque este resultado no puede generalizarse porque no siempre se encuentran de este modo (la figura 4, por ejemplo, presenta la correlación más baja en las posiciones finales).

Para responder a si alguno de los pares de métodos es más preciso en sus resultados de ordenación, en la tabla 3 se recoge el número de empates encontrados en los diferentes años y para las diferentes áreas de formación, con los distintos métodos.

Tabla 2
Matriz de correlaciones bivariadas entre métodos en función de los valores escalares (por encima de la diagonal) y la ordenación (por debajo de la diagonal)

	DR	OR	IS	IAI
DR	1	0.544	0.591	0.813
OR		1	0.417	0.793
IS			1	0.643
IAI				1

DR= método de Dunn-Rankin; OR= método de ordenación por rangos; IS= método de intervalos sucesivos; IAI= método de intervalos aparentemente iguales

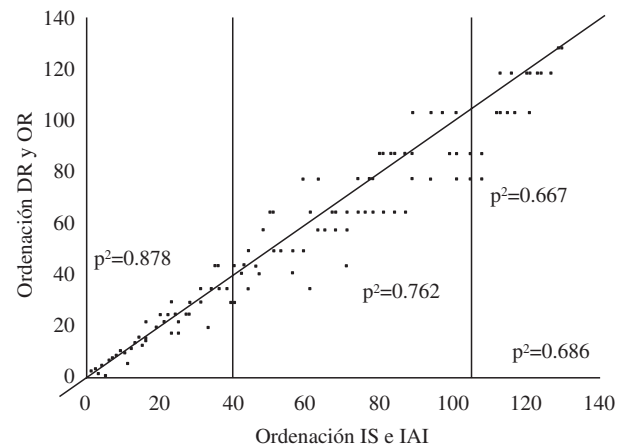


Figura 2. Diferencias en las ordenaciones en función de los métodos utilizados en el área de formación C, año 2006

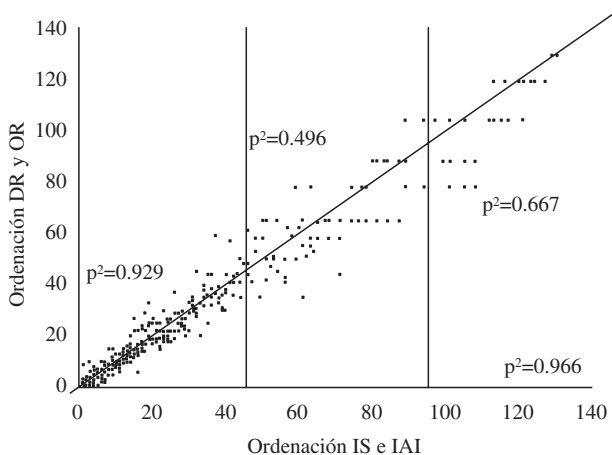


Figura 1. Diferencias en las ordenaciones en función de los métodos utilizados

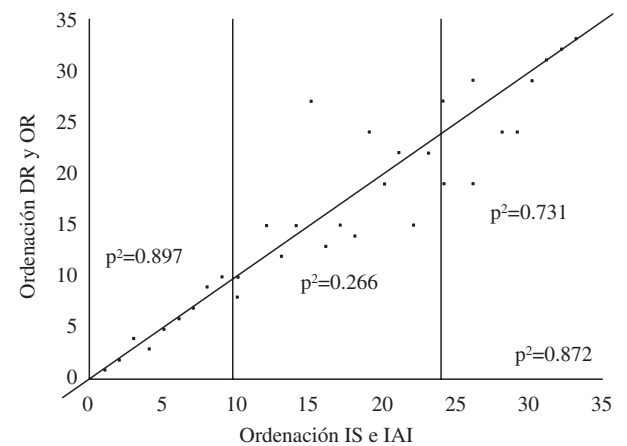


Figura 3. Diferencias en las ordenaciones en función de los métodos utilizados en el área de formación F, año 2006

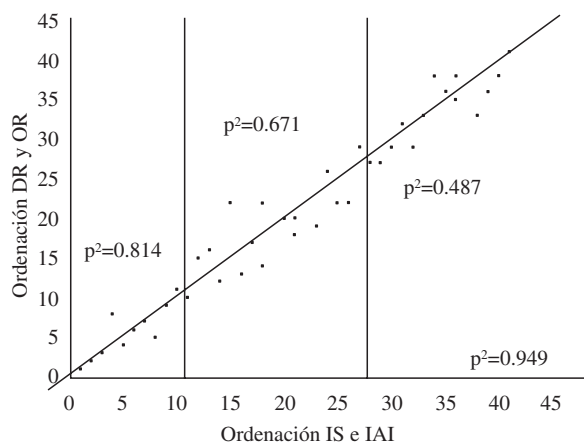


Figura 4. Diferencias en las ordenaciones en función de los métodos utilizados en el área de formación G, año 2006

Tabla 3						
Número de empates obtenidos en cada área y cada año con los diferentes métodos						
MET.	D-R/OR			IS/IAI		
AÑO	04	05	06	04	05	06
A	1	0	0	1	0	1
B	1	0	0	1	1	3
C	8	2	57	15	8	109
D	-	-	0	-	-	0
E	0	0	0	0	0	0
F	0	0	3	0	0	10
G	3	2	3	6	2	11
H	0	-	0	0	-	0
I	1	0	0	0	0	1
J	-	0	0	-	0	0
K	0	0	0	2	0	1
TOTAL	81			172		

DR= método de Dunn-Rankin; OR= método de ordenación por rangos; IS= método de intervalos sucesivos; IAI= método de intervalos aparentemente iguales

En líneas generales, destaca el elevado número de empates en el área C, seguida a distancia del área G. Atendiendo a las diferencias entre métodos, se constata que el número de veces que un estímulo comparte su posición con otro es mayor aplicando los métodos de intervalos sucesivos y aparentemente iguales (172), que con los métodos de Dunn-Rankin y ordenación por rangos (81), constatándose con la prueba de rangos señalados y pares igualados de Wilcoxon que esta diferencia es estadísticamente significativa con un nivel de confianza del 99% ($z = -2.99$; $p = 0.003$), por lo que puede concluirse que los métodos de Dunn-Rankin y de ordenación por rangos son los que presentan mayor precisión en sus resultados.

Discusión y conclusiones

La implantación de los métodos de escalamiento como procedimiento sistemático y transparente de procesamiento de informa-

ción aporta una solución empírica a los problemas existentes en la evaluación de necesidades formativas: permite explicitar los criterios para priorizar necesidades formativas a partir de la pertinencia establecida desde los implicados en el proceso; delimita criterios de asignación de recursos a acciones; y aporta criterios de inclusión/exclusión de los usuarios al programa.

De este modo, constatamos la factibilidad de estos métodos en el contexto de la formación continua en las organizaciones, porque pudo llevarse a cabo sin generar por ello discrepancia alguna con el funcionamiento cotidiano de la organización; y su utilidad porque se obtuvo, a partir de la participación activa de los implicados, un amplio número de acciones formativas ordenadas en función de su grado de relevancia, que sirvió como criterio de toma de decisiones para la conformación del plan de formación.

En relación al estudio de la precisión de los resultados obtenidos con los diferentes métodos, en función de las correlaciones bivariadas, podría concluirse que no existen grandes diferencias en las ordenaciones, resultado acorde con la literatura. El hecho de que se encuentren en ocasiones más diferencias en los estímulos escalados en las posiciones centrales puede deberse a que la ordenación dada por los sujetos en los estímulos extremos sea más homogénea al tratarse de estímulos más salientes, considerándose claramente acciones formativas muy necesarias o no necesarias en absoluto mientras que, en los estímulos centrales, las respuestas pudieron ser más heterogéneas, obteniéndose en consecuencia valores escalares más cercanos entre sí.

Aunque este resultado podría interpretarse como una debilidad de los métodos evidenciada cuando el número de estímulos a escalar es amplio, algunos indicios descartan esta explicación como posibilidad: (a) un número más elevado de estímulos a escalar no implica una correlación más baja en la ordenación obtenida con los distintos métodos en los estímulos centrales frente a los extremos, tal y como se puede comprobar en las figuras 2, 3 y 4 presentadas; (b) por el modelo teórico seguido descrito previamente en el apartado de procedimiento, los participantes en la ordenación de estímulos previamente colaboraron en la elaboración del listado de estímulos a priorizar, por lo que se encontraban familiarizados con el contenido de dichos estímulos y la posible tasa de error debida a la exposición de un número elevado de estímulos novedosos quedaría neutralizada; y (c) en estudios de años posteriores se solicitó a los participantes el escalamiento, dentro de cada área, de los 10 estímulos que consideraran más relevantes; a pesar de la disminución en la dificultad de la ordenación a realizar, el grado de acuerdo entre métodos seguía siendo menor en los estímulos centrales de manera mayoritaria.

Sin embargo, en función del número de empates obtenidos (de ordenaciones iguales para diferentes estímulos), los métodos de Dunn-Rankin y de ordenación por rangos son más precisos por encontrarse en éstos un número significativamente menor de tales empates.

Respecto al grado de sencillez en su procedimiento se concluye, de acuerdo con la literatura, que el método de los intervalos sucesivos es el más complejo por dos razones: (a) presenta la limitación de imposibilidad del cálculo del valor escalar para los estímulos situados en los extremos, por lo que se hace necesario un método complementario para obtener la ordenación en estos casos; y (b) implica el cálculo de diferentes amplitudes y límites para cada intervalo (Cuesta et al., 1997); le siguen de mayor a menor nivel de complejidad, el método de intervalos aparentemente iguales, el método de Dunn-Rankin y, finalmente, como más sencillo, el método de ordenación por rangos.

En síntesis, a pesar de que a grandes rasgos apenas existen diferencias en los resultados en función del método utilizado, se recomienda el uso del método de ordenación por rangos, pues conlleva la mayor precisión en sus resultados con el mayor grado de sencillez en su procedimiento; seguido del método de Dunn-Rankin, igual de preciso que el anterior, pero algo menos sencillo en su procedimiento; posteriormente, el método de intervalos aparentemente iguales; y, finalmente, el método de intervalos sucesivos por resultar el menos preciso, más complejo y con más limitaciones en su uso.

Consideramos que este trabajo aporta solución a un problema no baladí en ciencias sociales. El gran hándicap al que se enfrenta la intervención social y educativa es la falta de sistematización sobre las variables intervenidas, lo que dificulta enormemente el avance acumulativo de su conocimiento científico. La identificación operativa de los constructos estudiados posibilita la validación de modelos teóricos referentes.

En estudios futuros, complementariamente, además de los cuestionarios propuestos, se aplicarán instrumentos de formato abierto como son los grupos de discusión y entrevistas, para trian-

gular los distintos resultados obtenidos y comprobar el nivel de concordancia entre métodos.

Este trabajo posibilita además la informatización del procedimiento, la recogida de datos vía internet y la emisión de informes *in situ* a lo largo del proceso. Esto potencia la delimitación transparente de criterios de valoración, participación, utilidad y rigor metodológico.

Quedamos a disposición para aportar información adicional y para colaboraciones en este contexto y/o la extrapolación a distintos ámbitos.

Agradecimientos

El presente trabajo forma parte de los resultados obtenidos en los proyectos de investigación SEJ 2004-05360/EDUC, subvencionado por el Ministerio de Educación y Ciencia; y P06-HUM-01458, subvencionado por la Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía.

Los datos se obtuvieron a través del convenio de colaboración renovado anualmente entre la Diputación de Sevilla y la Universidad de Sevilla (SI-39/04, SI-87/05 y SI-114/06).

Referencias

- Anguera, M.T., Chacón, S., y Blanco, A. (2008). *Evaluación de programas sociales y sanitarios. Un abordaje metodológico*. Madrid: Síntesis.
- Arce, C. (1994). *Técnicas de construcción de escalas psicológicas*. Madrid: Síntesis.
- Arce, C., y Andrade, E.M. (2000). Recuperación de información métrica a partir de información no-métrica en diseños de escalamiento multidimensional incompletos. *Psicothema*, 12(2), 308-313.
- Barbero, M.I. (1993). *Psicometría II. Métodos de elaboración de escalas*. Madrid: UNED.
- Chacón, S., Anguera, M.T., Pérez-Gil, J.A., y Holgado, F.P. (2002). A mutual catalytic model of formative evaluation: The interdependent roles of evaluators and local practitioners. *Evaluation. The International Journal of Theory, Research and Practice*, 8(4), 413-432.
- Chacón, S., Lara, A., y Pérez-Gil, J.A. (2002). Needs assessment. En Fernández-Ballesteros, R. (Ed.): *Encyclopedia of psychological assessment*, vol. 2 (pp. 615-619). London: Sage.
- Chacón, S., López, J.M., y Sanduvete, S. (2004). *Evaluación de acciones formativas en Diputación de Sevilla*. Una guía práctica. Sevilla: Diputación de Sevilla.
- Coombs, C.H. (1964). *A theory of data*. New York: Wiley.
- Cuesta, M., Jiménez, J.M., Álvarez, B., Díaz, R., y Paz, D. (1997). Análisis de las propiedades psicométricas de una escala de actitud: comparación de las técnicas de Likert y Thurstone. *Revista Electrónica de Metodología Aplicada (REMA)*, 2(2), 23-33.
- Gómez, A.M., y Palací, F.J. (2003). El constructo «Orientación a la calidad total» y su medida: el Cuestionario OCT. *Psicothema*, 15(3), 433-439.
- Little, R.J.A., y Rubin, D.B. (2002). *Statistical analysis with missing data*. New York: Wiley.
- Meliá, J.L., Sospedra, M.J., Molina, J.G., Ramón, E.G., e Islas, M.E. (1991). Medición de objetos psicológicos: ajuste lineal entre las escalas obtenidas por cuatro métodos de escalamiento unidimensional. *Psicológica*, 12(1), 35-45.
- Meliá, J.L., Sospedra, M.J., Ramón, E.G., y Molina, J.G. (1992). Escalamiento de la gravedad de delitos contra la mujer: valoración legal versus valoración social. *Anuario de Psicología*, 52, 111-120.
- Ponsoda, V. (1986). *Iniciación a la psicología matemática*. Madrid: Universidad Autónoma.
- Sanduvete, S. (2008). *Innovaciones metodológicas en la evaluación de la formación continua*. Tesis doctoral no publicada: Universidad de Sevilla.
- Siegel, S. (1991). *Estadística no paramétrica aplicada a las ciencias de la conducta*. México D. F.: Trillas.
- Vedung, E. (1995). Utilización de la evaluación de programas. En R. Fernández-Ballesteros (Ed.): *Evaluación de programas. Una guía práctica en ámbitos sociales, educativos y de salud* (pp. 321-338). Madrid: Síntesis.
- Wen, K., y Gustafson, D.H. (2004). Needs assessment for cancer patients and their families. *Health and Quality of Life Outcomes*, 2:11.
- White, C., Turoff, M., y Van de Walle, B. (2007). A dynamic Delphi process utilizing a unified Thurstone scaling method: Collaborative judgment in emergency response. En P. Burghardt, K. Nieuwenhuis y B.A. van de Walle (Eds.): *Proceedings of the 4th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2007)* (pp. 7-15). The Netherlands, Delft.
- Witkin, B.R., y Altschuld, J.W. (1995). *Planning and conducting needs assessments. A practical guide*. London: Sage.