

La matriz multimétodo-multirrasgo aplicada al estudio de la sensibilidad

Pedro Rodríguez-Miñón Cifuentes, Enrique Moreno González y Pilar Sanjuán Suárez
Universidad Nacional de Educación a Distancia

El escalamiento psicofísico dispone de distintas medidas de la sensibilidad que han sido clasificadas en diferentes categorías, entre las que cabe señalar aquí a las provenientes del modelo de Stevens. Estudios previos (Villarino, Garriga-Trillo y Rodríguez-Miñón, 1992; Villarino, Garriga-Trillo, 1995) han comparado estas medidas de la sensibilidad, encontrando altas correlaciones entre las derivadas del modelo de Stevens, pero bajas cuando se comparan con las provenientes de otros modelos (los de la Teoría de Detección de Señales y los de la Teoría de la Información), lo que puede sugerir que estas diferentes medidas no están midiendo lo mismo. El objetivo del presente trabajo es abordar, en el campo del escalamiento psicofísico, el estudio de la fiabilidad y validez del exponente de Stevens como medida de la sensibilidad, a partir de la matriz multimétodo-multirrasgo. Aunque dicho procedimiento se ha utilizado para la validación de pruebas psicológicas, nunca se ha aplicado para la validación de las medidas de la sensibilidad derivadas del campo de la psicofísica. En esta área de investigación, la aplicación de este procedimiento conlleva la evaluación de diversas modalidades sensoriales (visión, audición, tacto, etc.) o, en nuestro caso, estimaciones de distintas dimensiones (longitud, superficie, luminosidad, etc.) mediante distintos procedimientos de obtención de datos (estimación de magnitudes absoluta y estimación de magnitudes con diferentes estímulos estándares).

The multimethod-multitrait matrix applied to the study of sensitivity. The psychophysical scaling has several sensitivity measurements that have been classified in different groups as, i.e. those that derived of the Stevens model. Previous researchs (Villarino, Garriga-Trillo & Rodríguez-Miñón, 1992; Villarino & Garriga-Trillo, 1995) have compared these measurement of sensitivity and have found high correlations in the measures derived of Stevens model, but low correlations when another models are considered (Signal Detection Theory and Information Theory). These results could suggest that these measurements do not measure the same. The main objective of present work is to approach study of reliability and validity of Stevens's exponent as measurement of sensitivity through multimethod-multitrait matrix. Though this procedure has been used for validation of psychological tests, it never has been applied in the psychophysical field. In this investigation area, the application of this procedure involve the evaluation of different sensory kind (view-sight, hearing, touch-tact, etc.) or, in our case, estimation of several dimensions (length, area, brightness, etc.) through different procedures of data obtaining (absolute magnitude estimation, and magnitude estimation with several standard stimuli).

Las técnicas de escalamiento psicofísico provenientes del modelo de Stevens se apoyan en la hipótesis de que los sujetos puedan realizar juicios de la intensidad percibida de los estímulos asignando números en proporción directa a las sensaciones suscitadas por dichos estímulos. Con la aplicación de estas técnicas de escalamiento, el exponente de la función potencial de Stevens se comporta, en muchos casos, como una medida de la sensibilidad (Villarino, Garriga-Trillo, Rodríguez-Miñón, García-Gallego y Merino, 1991) pero en otras ocasiones los resultados son contrarios (Villarino, 1994) al depender de las condiciones experimenta-

les, de la modalidad sensorial y, especialmente, de la utilización o no de un estímulo estándar. Cuando se utiliza un estímulo estándar, el exponente correlaciona positiva y significativamente con otras medidas de la sensibilidad, pero cuando no se utiliza el estándar estas correlaciones no aparecen (Villarino y Garriga-Trillo, 1995) lo que indicaría que el uso del estándar mejora el exponente de Stevens como medida de la sensibilidad al favorecer la consistencia de las respuestas de los sujetos. Igualmente, en los ajustes individuales, Coello, Rodríguez-Miñón, Fontes, Merino y Garriga-Trillo, (1992) encuentran que las diferencias individuales en los exponentes dependen de la modalidad con la que se trabaje siendo mayor para la modalidad visual y mucho menor para la táctil, interpretando estos resultados en el sentido de que la modalidad táctil es un sentido más burdo y por tanto más difícil de que aparezcan estas diferencias, máxime si la escala física está más acotada lo que conlleva que los juicios emitidos se asemejen más entre ellos. Otros autores sugieren que la variabilidad en los expo-

nentes individuales se debe a la forma peculiar que los sujetos tienen de asignar los números a las sensaciones, y no a las diferencias en los procesos sensoriales (Algorn y Marks, 1984; Hellman, 1981).

Otra fuente de variabilidad puede atribuirse a los sesgos de respuesta en los juicios psicofísicos. Poulton (1989) es el autor que más extensamente ha estudiado estos sesgos de respuesta en los juicios cuantitativos de las magnitudes sensoriales entre los que cabe citar, por su posible aparición en este experimento, la posición del estímulo estándar, su módulo y los efectos de contexto. Cuando se utiliza un estándar situado en los extremos del rango se producen exponentes menores que cuando se utiliza un estándar situado en el rango medio estimular (Engen y Ross, 1966). Otros autores (Rodríguez-Miñón, Garriga-Trillo, García-Gallego y Merino, 1990) encuentran, con longitudes de segmentos que, para mejorar la estimación, el estímulo estándar debe ser un valor próximo al comparativo y para aproximar el exponente empírico al valor teórico aceptado (la unidad con longitudes de segmentos) el estímulo estándar debe ser un valor extremo del rango estimular. También el módulo asignado al estándar produce sesgos de respuesta, de tal forma que, cuando se incrementa el valor del módulo el valor del exponente disminuye (Wong, 1963). Estudiando los efectos secuenciales para diferentes estímulos comparativos utilizando el mismo módulo asignado al estándar se encuentra que las respuestas vienen dadas en función de este módulo, con independencia de la magnitud del estímulo a juzgar (García-Gallego, Garriga-Trillo, Merino, Rodríguez-Miñón y Villarino, 1991; Villarino; Garriga-Trillo y Rodríguez-Miñón, 1992). Los efectos de contextos utilizando diferentes técnicas de escalamiento numérico han sido encontrados por algunos autores (Marks, 1992, 1993; García-Gallego y Garriga-Trillo, 1995). Ward (1990) sugiere que intervienen tanto los procesos perceptuales como los sensoriales en estos efectos contextuales. Todas estas fuentes de error en la aplicación de la estimación de magnitudes, llevaron a Zwislocki (1989) a plantear una alternativa, conocida como estimación de magnitudes absoluta (EMA), que se apoya en la idea de que los sujetos pueden ser instruidos para emparejar su impresión subjetiva del tamaño de un número a su impresión de la intensidad de un estímulo, produciendo una escala psicofísica con una unidad fija natural y que es insensible a los efectos de contexto estimular (Gescheider, 1990). De acuerdo con la clasificación de las escalas de medida dadas por Stevens, una escala absoluta consistiría en una escala de razón con una unidad fija, y contendría todas las propiedades de los otros tipos de escala (nominal, ordinal, intervalo y razón). Cuando, por el contrario, con la estimación de magnitudes convencional se fuerza al observador a utilizar un módulo arbitrario asociado con un estímulo estándar particular, para forzar que todos los sujetos utilicen la misma unidad de medida, aparecen los sesgos de respuesta (Gescheider y McDonnell, 1991) de los que una pequeña muestra se han mencionado en líneas anteriores.

En cualquier caso, si el exponente de Stevens es una medida de la sensibilidad, debe poseer las mismas propiedades que cualquier otra medida, a saber: fiabilidad y validez. En una investigación anterior (Rodríguez-Miñón, Garriga-Trillo y Villarino, 1993), mediante el método de test-retest, se abordó el estudio de la fiabilidad de los juicios psicofísicos de la longitud de segmentos, obteniéndose coeficientes superiores a 0.91, lo que indica que los sujetos tienden a repetir las respuestas emitidas a los mismos estímulos en presentaciones realizadas en días sucesivos.

En el presente trabajo pretendemos estudiar dichas características mediante la matriz multimétodo-multirrasgo (Campbell y Fiske, 1959). Este procedimiento se ha utilizado ampliamente para la validación de pruebas psicológicas, pero nunca se ha aplicado para la validación de las medidas de la sensibilidad derivadas del campo de la psicofísica. En este área, la aplicación de este procedimiento implica la evaluación de diversas modalidades sensoriales, como visión, audición, tacto, etc. mediante distintos procedimientos de obtención de datos (por ej., estimación de magnitudes, producción de magnitudes, etc.). En concreto, en la presente investigación utilizamos las dimensiones de longitud (radios y diámetros) y superficie (círculo y sector circular) valoradas mediante el procedimiento de estimación de magnitudes, en sus variantes de estimación de magnitudes absoluta y estimación de magnitudes con diferentes estímulos estándares.

En la matriz multimétodo-multirrasgo, la fiabilidad está representada por la intensidad del acuerdo de los exponentes (la sensibilidad) obtenidos al escalar el mismo estímulo a partir de técnicas idénticas de escalamiento. Por otra parte, se puede hablar de validez convergente cuando hay correlaciones elevadas, y significativamente distintas de cero, utilizando los exponentes de un mismo estímulo valorado con distintas técnicas y estas correlaciones deben ser más altas que las correlaciones entre los exponentes obtenidos con diferentes estímulos escalados con el mismo método.

Método

Sujetos

La muestra estaba formada por 12 sujetos, 6 mujeres y 6 varones, que participaron voluntariamente en el experimento y pasaron por todas las condiciones experimentales

Procedimiento

Como técnica de escalamiento se utilizó la estimación de magnitudes en todas las condiciones. Esta técnica tiene la enorme ventaja de poder registrar gran cantidad de estimaciones en un periodo razonablemente corto de tiempo. La técnica se utilizó en tres versiones diferentes: 1) Estimación absoluta de magnitudes (EMA), es decir, sin utilizar un estímulo estándar; 2) Estimación de magnitudes convencional (EMC) utilizando un estímulo estándar de magnitud intermedia; y 3) Estimación de magnitudes convencional (EMC) utilizando un estímulo estándar de magnitud máxima.

Se utilizaron cuatro tipos de estímulos: 1) Radio del círculo; 2) Diámetro del círculo; 3) Área del círculo; y 4) Área del sector circular (que en todos los casos era el 58% del círculo del que provenía). Cada estímulo se presentó con 20 magnitudes diferentes y 7 veces cada uno de ellos, por lo que cada sujeto juzgaba un total de 140 estímulos presentados en orden aleatorio. Las magnitudes de los estímulos eran proporcionales, es decir, el estímulo número 20 –el de mayor tamaño– era 20 veces la magnitud del estímulo menor, en los estímulos de longitud (radio y diámetro) y 400 veces mayor en los estímulos de superficie (área del círculo y del sector circular). Por tanto, las razones entre las diferentes magnitudes de los estímulos de superficie eran iguales al cuadrado de la razón de esas magnitudes en los estímulos de longitud. En cada una de las sesiones, los sujetos realizaban tantos ensayos de prueba como deseaban, para pasar posteriormente a realizar las 140 estimaciones de la condición concreta que estuvieran realizando.

Para paliar el posible efecto del orden de realización de las tareas sobre las estimaciones, un tercio de los sujetos (4; 2 varones y 2 mujeres) comenzó con la EMA, otro tercio por la EMC con estímulo estándar intermedio y el tercio restante con la EMC con estímulo estándar extremo. Ya dentro de cada uno de estos grupos y tareas se impuso la restricción de que los observadores no podían realizar estimaciones consecutivas de objetos geométricos con las mismas características, es decir, no podían estimar radios y a continuación diámetros, sino que se intercalaban con las estimaciones de áreas (círculos o sectores)

Resultados

Para cada sujeto se calculó el exponente, «n», de la función de Stevens o, lo que es lo mismo, la pendiente, b, de la ecuación de regresión del logaritmo neperiano de la respuesta sobre el logaritmo neperiano del estímulo. Este parámetro se obtiene para cada técnica de estimación y con cada uno de los estímulos utilizados, con lo que disponemos de 144 exponentes (12 sujetos x 12 condiciones experimentales). Posteriormente, y a partir de estos exponentes se obtiene la matriz de correlaciones, o matriz multimétodo-multirasgo, que puede verse en la tabla I.

Siguiendo a Campbell y Fiske (1959) en la diagonal principal de la matriz se encuentran las correlaciones calculadas entre los exponentes obtenidos con idénticos estímulos escalados con la misma técnica de estimación (correlaciones monométodos-monorrasgo) y que representan el coeficiente de fiabilidad —obtenida mediante el procedimiento de división en mitades y aplicando posteriormente la corrección de Spearman-Brown—. Por la parte inferior de esta diagonal principal, también podemos observar tres triángulos contiguos que incluyen las correlaciones entre los exponentes obtenidos con diferentes estímulos escalados con la misma técnica de estimación (correlaciones monométodo-heterorrasgo). Los bloques heterométodo están formados por una diagonal de correlaciones de los «n» obtenidos para los mismos estímulos escalados con distintas técnicas de estimación (correlaciones heterométodo-monorrasgo) y dos triángulos de correlaciones heterométodo-heterorrasgo que no son simétricos. Para que exista validez, los valores de la diagonal deben ser grandes y significativamente distintos de cero, indicando una convergencia entre los ex-

ponente obtenidos para los mismos estímulos escalados con distintas técnicas de estimación y mayores que las correlaciones entre los exponentes heterorrasgo-heterométodo, que son las calculadas entre exponentes que no tienen en común ni el estímulo ni la técnica de estimación.

Discusión

A la vista de nuestros resultados podemos afirmar que con las tres técnicas de estimación se obtienen altos coeficientes de fiabilidad, lo que vienen a confirmar los obtenidos con los juicios psicofísicos estudiados anteriormente utilizando métodos de cálculo de la fiabilidad dirigidas al estudio de la estabilidad temporal (Rodríguez-Miñón, Garriga-Trillo y Villarino, 1993). Los resultados obtenidos indicarían una consistencia en los exponentes individuales, aunque desde un punto de vista heurístico, parece que la EMA es la que conduce a una mayor fiabilidad. La estimación de magnitudes con un estándar situado en el extremo superior del rango estimular es la que produce coeficientes de fiabilidad inferiores, obteniéndose, para la estimación del radio, una correlación negativa que es, de acuerdo con el modelo lineal planteado por Spearman, un valor imposible.

Por lo que respecta a los bloques heterométodo-monorrasgo, ninguno de los intentos de medir el mismo estímulo con diferentes técnicas conduce a resultados consistentes lo que parece indicar que el exponente de Stevens carece de validez convergente como medida de la sensibilidad. Si además, observamos las correlaciones significativas que se producen al escalar diferentes estímulos con la misma técnica, EMA, podemos inferir que los exponentes obtenidos dependen de la técnica empleada para recoger los juicios psicofísicos, es decir, en los ajustes individuales se obtienen para cada sujeto y para cada tipo de estímulos exponentes más similares al utilizar la EMA que cuando se utiliza la EMC. Obsérvese, por otra parte, que todos los estímulos están relacionados entre ellos: en dos hay que estimar longitudes: radio y diámetro y en los otros dos hay que estimar áreas, que dependen de las longitudes del radio y/o del diámetro. Aplicar la misma técnica para estimar diferentes estímulos, con una característica común que los relaciona, conduce a altos coeficientes de correlación cuando se utiliza la EMA que con la

Tabla 1

	ESTIMACIÓN ABSOLUTA				E. RELATIVA (ST= 20)				E. RELATIVA (ST= 5)			
	C	S	D	R	C	S	D	R	C	S	D	R
C	0.986											
S	0.67 *	0.966										
D	0.52	0.84 **	0.969									
R	0.85 **	0.85 **	0.71 **	0.881								
C	0.058	0.304	0.21	0.036	0.562							
S	0.26	0.032	-0.034	0.063	0.32	0.772						
D	-0.026	0.40	0.087	0.10	0.22	0.134	0.711					
R	0.059	-0.098	0.026	-0.074	0.21	0.17	-0.57	-0.347				
C	0.14	-0.096	-0.19	0.006	-0.446	0.092	0.189	-0.505	0.941			
S	0.11	0.045	0.046	0.157	0.04	0.252	-0.070	-0.25	0.442	0.918		
D	0.48	0.66	0.496	0.516	-0.264	-0.047	0.534	-0.539	0.431	0.084	0.932	
R	0.052	0.26	0.119	0.33	-0.259	0.271	0.455	-0.343	0.49	0.477	0.395	0.492

C=área del círculo
S=área del sector circular
D=diámetro
R=radio

EMC. Una posible explicación, a la luz de las investigaciones citadas previamente, puede deberse a que la EMA es más insensible a ciertos sesgos de respuesta que la EMC, lo que afecta a la validez del exponente de Stevens que pasaría, por tanto, a depender más de factores no sensoriales que de repuestas propiamente sensoriales. Con el fin de seguir refinando estas técnicas

de escalamiento psicofísico, los resultados obtenidos nos animan a seguir investigando en esta línea utilizando muestras más grandes, que aseguren mayor potencia en los contrastes, y diseñando experimentos que eviten en lo posible, o controlen en su defecto, los sesgos de respuesta más habituales en las técnicas de escalamiento psicofísico.

Referencias

- Algom, D. y Marks, L.E. (1984). Individual differences in loudness processing and loudness scales. *Journal of Experimental Psychology: General*, 113, 571-593.
- Campbell, D.T. y Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Coello, M.T., Rodríguez-Miñón, P., Fontes, S., Merino, J.M. y Garriga-Trillo, A.J. (1992). Diferencias individuales en la función potencial de Stevens: Un estudio entre modalidades. *Investigaciones Psicológicas*, 11, 153-163.
- Engen, T. y Ross, B.M. (1966). Effect of reference number on magnitude estimation. *Perception and Psychophysics*, 1, 74-76.
- García-Gallego, C. y Garriga-Trillo, A. (1995): Assimiling and contrast effects in length estimation task. En C.A. Possamai (Ed.), *Fechner Day 95*. Cassis: International Society for Psychophysics. (Pp. 191-194).
- García-Gallego, C., Garriga-Trillo, A.J., Merino, J.M., Rodríguez-Miñón, P. y Villarino, A. (1991). Sequential modelization of psychophysical data. En G. Lockhead (Ed.), *Fechner Day 91*. Durham: International Society for Psychophysics. (Pp. 49-54).
- Gescheider, G.A. (1990). Stimulus context and absolute magnitude estimation. En F. Müller (Ed.), *Fechner Day 90*. Würzburg: International Society for Psychophysics. (Pp. 43-48).
- Gescheider, G.A. y McDonnell, K.A. (1991). The effects of context on magnitude estimation and absolute magnitude estimation. En G. Lockhead (Ed.), *Fechner Day 91*. Durham: International Society for Psychophysics. Duke University. (Pp. 109-113).
- Hellman, R.P. (1981). Stability of individual loudness functions obtained by magnitude estimation and production. *Perception and Psychophysics*, 29, 63-70.
- Marks, L.E. (1992). The slippery context effect in psychophysics: Intensive, extensive, and qualitative continua. *Perception and Psychophysics*, 51 (2), 187-198.
- Marks, L.E. (1993): Differential context effects within but not between modalities?. En A.J. Garriga-Trillo, P. Rodríguez-Miñón, C. García-Gallego, P. Lubin, J.M. Merino y A. Villarino (Eds.), *Fechner Day 93*. Palma de Mallorca: International Society for Psychophysics. UNED. (Pp 161-166).
- Moreno, E. (1994). *Análisis de los juicios psicofísicos con la técnica de estimación de magnitudes*. Tesis Doctoral sin publicar. Madrid: UNED.
- Poulton, E.C. (1989). *Bias in quantifying judgments*. London: L.E.A.
- Rodríguez-Miñón, P., Garriga-Trillo, A.J., García Gallego, C. y Merino, J.M. (1990): Bias in line length estimation. En F.Müller (Ed.), *Fechner Day 90*. Würzburg: International Society for Psychophysics. (Pp. 94-99).
- Rodríguez-Miñón, P., Garriga-Trillo, A.J. y Villarino, A. (1993): Studying reliability in psychophysical judgements by classical test theory measurements. En A.J. Garriga-Trillo, P. Rodríguez-Miñón, C. García-Gallego, P. Lubin, J.M. Merino y A. Villarino (Eds.), *Fechner Day 93*. Palma de Mallorca: International Society for Psychophysics. UNED. (Pp 183-185).
- Villarino, A. (1994): *Medida de la sensibilidad gustativa: Una aplicación para la discriminación de vinos*. Tesis doctoral sin publicar. Madrid: UNED.
- Villarino, A. y Garriga-Trillo, A.J. (1995). The exponent of Stevens' law as a measure of sensitivity in line length estimation. En C.A. Possamai (Ed.), *Fechner Day 95*. Cassis: International Society for Psychophysics. (Pp: 315-319).
- Villarino, A., Garriga-Trillo, A.J. y Rodríguez-Miñón, P. (1992). More about the standard's bias. En G. Borg y G. Neely (Eds.), *Fechner Day 92*. Estocolomo: International Society for Psychophysics (pp. 211-216).
- Villarino, A., Garriga-Trillo, A.J., Rodríguez-Miñón, P., García-Gallego, C. y Merino, J.M. (1991): Comparing indices using different models and modalities. En G. Lockhead (Ed.), *Fechner Day 91*. Durham: International Society for Psychophysics. Duke University. (Pp. 95-98).
- Ward, L.M. (1990): Critical bands and mixed-frequency scaling: Sequential dependencies, equal-loudness contours, and power function exponents. *Perception and psychophysics*, 46(6), 551-562.
- Wong, R. (1963). Effect of the modulus on estimates of magnitude of linear extent. *American Journal of Psychology*, 76, 511-512.
- Zwischlocki, J.J. (1989). Natural Measurement. En S.J. Bolanowski y G.A. Gescheider (Eds.), *Ratio scaling of psychological magnitude*. Hillsdale: L.E.A.