

“La Psicología como ciencia y su relación con otras disciplinas”

MEMORIAS EN EXTENSO

**XLIII Congreso del Consejo Nacional para la Enseñanza e Investigación en Psicología A.C.
(CNEIP)**

Xalapa, Veracruz del 27 al 29 de abril del 2016

**Universidad Veracruzana,
Facultad de Psicología**

Editado por:

Dra. Ana Gloria Gutiérrez-García
Mtro. Juan Grapain-Contreras
Mtro. Herson A. Castellanos-Celis

Facultad de Psicología, Universidad Veracruzana
Manantial de San Cristóbal s/n Xalapa, 2000
Xalapa, Veracruz. México

ISBN 978-607-97084-2-9

Citar de la siguiente manera:

Gutiérrez-García Ana G., Grapain-Contreras Juan, Castellanos-Celis Herson A. Compiladores. (2016). La Psicología como ciencia y su relación con otras disciplinas. XLIII Congreso del Consejo Nacional para la Enseñanza e Investigación en Psicología A.C. (CNEIP). México: Universidad Veracruzana.

Validación psicofísica de opciones en escalas tipo Likert de frecuencia.

Manuel Alejandro Mejía Ramírez

Universidad de las Californias Internacional, Tijuana, Baja California, manuel.mejia@udc.edu.mx

RESUMEN

Las escalas tipo Likert han sido ampliamente usadas en medición en psicología y en múltiples ámbitos. Permiten cuantificar múltiples tipos de comportamiento, de los más usados es el que cuantifica frecuencia (ej. casi siempre, a veces). El escalamiento que propuso Likert supone que las categorías de respuesta de cada ítem se interpretan a nivel de intervalos, asignando números a las respuestas. Sin embargo, no existen suficientes estudios que verifiquen la interpretación que las personas les dan a las opciones de respuesta. Si esto no se verifica, no se puede asumir una escala de intervalos para poder sumar las respuestas obtenidas, que usualmente se realiza. Cañadas y Sánchez (1998) atajaron este problema en España, sin embargo, las diferencias lingüísticas entre España y México, y su evolución lingüística, exigen realizar la validación en México. Establecer puntuaciones a nivel de intervalos de los diferentes cuantificadores lingüísticos de frecuencia que se utilizan en escalas tipo Likert en México mediante una técnica psicofísica. Se utilizó el método de estimación de magnitudes de Stevens con una lista de 27 cuantificadores lingüísticos de frecuencia. Con esta tarea se obtuvieron cuantificaciones de cada estímulo, por participante, en una escala relativa a dos estímulos de referencia (contrabalanceados). Los 27 estímulos utilizados fueron ordenados según la media geométrica obtenida de cada uno a partir de las calificaciones dadas por 89 participantes. Los resultados representan un orden lógico del significado de cada cuantificador lingüístico. La cuantificación intervalar obtenida de cada estímulo (la media geométrica) permite establecer una selección de opciones para escalas tipo Likert basadas en intervalos iguales según el número de opciones de interés. En el caso de 5 opciones, los resultados proponen una escala con: *nada/nunca, alguna vez, a menudo, casi siempre, y siempre*. Los resultados fundamentan la correcta interpretación intervalar de las opciones de respuesta.

1. INTRODUCCIÓN

Desde que se desarrollaron escalas Likert para medir actitudes, éste método ha sido utilizado en escenarios diversos, como mercadotecnia, enfermería, entre otros (Joshi, Kale, Chandel, & Pal, 2015). La técnica desarrollada por Likert (1932) ha demostrado su utilidad al permitir el uso de cuantificadores lingüísticos de diverso tipo, uno de los más usuales es el que identifica frecuencia (ej. “casi siempre”, “a veces”, “frecuentemente”).

Al elaborar una escala tipo Likert, la selección de cuáles cuantificadores lingüísticos de frecuencia se usarán ha sido realizada comúnmente bajo criterios variados. El criterio que parece predominar en la práctica es el de la intuición del constructor del instrumento.

Además del problema de la selección de las opciones de respuesta (cuantificadores), el nivel de medición que representan ha sido un tema de debate, siendo una de las opiniones estrictas el que éstas escalas únicamente alcanzan el nivel ordinal (Boone & Boone, 2012; Gadermann, Guhn, & Zumbo, 2012; Li, 2013). Si las escalas tipo Likert son de nivel ordinal, no se puede justificar la suma o promedio de los

puntajes obtenidos en los reactivos, que es lo que comúnmente se realiza para su análisis.

Por la razón anterior, las mediciones obtenidas con escalas tipo Likert, de igual manera, tendrían la problemática de no justificar el uso de herramientas estadísticas paramétricas para su análisis. A pesar de que ha habido propuestas de solución argumentando la robustez de las pruebas estadísticas paramétricas aún en el análisis de datos ordinales (Boone & Boone, 2012; Norman, 2010), no se resuelve la situación de que éstas escalas no cumplen el nivel de medición de intervalos.

Por lo tanto, se hace necesario el aplicar un método que permita verificar el nivel de medición que se obtiene con las opciones de respuesta en una escala tipo Likert. El método permitiría, entonces, el realizar una selección objetiva de las opciones de respuesta al construir una escala, dando mayor certeza en que el nivel de medición obtenido sea de intervalos.

Una propuesta desarrollada en Estados Unidos (Schriesheim & Castro, 1996; Schriesheim & Gardiner, 1992; Schriesheim & Schriesheim, 1978) y España (Cañadas Osinski, Prieto Marañón, San Luis Costas, & Caballero de Rodas, 1994; Cañadas Osinski & Sánchez Bruno, 1998, 2000) para estimar un nivel de intervalos al elegir opciones de respuesta en una escala tipo Likert ha sido usando el método de estimación de magnitudes desarrollado por Stevens (Fontes, Fontes, & Merino, 1990; Stevens, 1956) en el ámbito de la psicofísica.

Sin embargo, el uso de esta técnica requiere que se realice la validación en el contexto lingüístico donde se utilizarán las escalas construidas (Cañadas Osinski & Sánchez Bruno, 1998), así como el realizar una verificación transcultural de la utilidad del mismo método con estímulos lingüísticos. El presente trabajo hace uso del método de estimación de magnitudes en población mexicana universitaria, y permite verificar su utilidad para validar el nivel de intervalos en escalas tipo Likert, así como establecer recomendaciones de cuáles opciones de respuesta usar en su construcción.

2. MÉTODO

Se utilizó el método de estimación de magnitudes de Stevens (1956) con un total de 27 cuantificadores lingüísticos de frecuencia. Los cuantificadores se obtuvieron del estudio de Cañadas Osinski y Sánchez Bruno (1998) y fueron complementados a partir de un sondeo de diferentes escalas tipo Likert publicadas en México.

Instrumentos

Los cuantificadores lingüísticos se presentaron en computadora, a través del programa OpenSesame (Mathôt, Schreij, & Theeuwes, 2012), que permitió la automatización en la recolección de datos. Los participantes respondieron usando el teclado de la computadora.

Participantes

Participaron 88 estudiantes de varias licenciaturas, de tres universidades diferentes de la ciudad de Tijuana, de entre 18 y 50 años, de ambos sexos.

Procedimiento

Se convocó a los participantes directamente en los salones de clases, y se les condujo al laboratorio de cómputo de la universidad. Los participantes se sentaron frente a la computadora, donde se obtuvo el consentimiento informado, y se dieron instrucciones.

Se presentó uno de los estímulos referentes y se pidió al participante asignar un número a éste, de manera que representara subjetivamente su valor en frecuencia. Enseguida se presentó el resto de los estímulos en orden aleatorio para cada participante, y el participante debía asignar un número a cada uno, de manera que fuera relativo a lo asignado al primer estímulo (ver Figura 1).

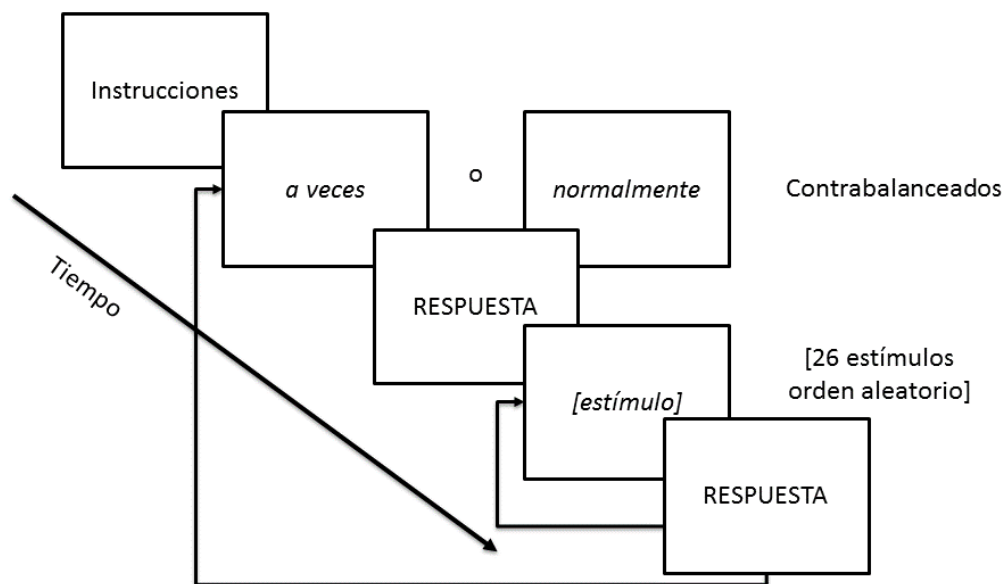


Figura 1. Estructura de la sesión experimental. En un primer momento los participantes leyeron las instrucciones, luego se les presentó un primer referente, al que tenían que calificar escribiendo una respuesta en la pantalla siguiente. Después de esto, se les presentaron el resto de los 26 estímulos en orden aleatorio y tras cada uno de ellos, se capturaba su respuesta. Después se repitió el procedimiento cambiando el referente inicial.

Los estímulos usados como referentes fueron “a veces” y “normalmente”. La mitad de los participantes usó primero el referente “a veces”, y la otra mitad usó en primer lugar “normalmente”. Después de realizado esto, se aplica la misma dinámica, usando el otro estímulo referente, para comprobar los resultados. Se realiza un análisis estadístico a partir de las medias geométricas de los números asignados a cada estímulo por todos los participantes.

Para comprobar la consistencia de los resultados con los dos estímulos usados como referentes (“a veces” y “normalmente”), se obtuvo la correlación entre las medias geométricas calculadas para cada estímulo con cada referente. Adicionalmente, se compararon usando una *t* de Student para muestras relacionadas.

Para obtener valores promedio a partir de las cantidades que los participantes asignaron a los estímulos se utilizó la media geométrica, esta herramienta nos permite analizar los resultados puesto que cada participante podía usar la escala que subjetivamente le pareciera mejor. Para poder incluir valores iguales a cero, se sumó a cada valor dado por los participantes la constante de 0.01. Después de esta suma, se

obtuvo el logaritmo base 10 de todos los valores, y se calculó la media aritmética de estos logaritmos para cada uno de los 27 estímulos por separado. De cada media aritmética se calculó su antilogaritmo, y se les restó el valor 0.01. Este valor así obtenido por cada uno de los estímulos, se consideró la media geométrica, reportada en la última columna de la Tabla 1.

3. RESULTADOS

A partir de los resultados obtenidos de los participantes usando como referentes dos estímulos, se filtraron los datos de aquellos participantes que hubieran asignado valores que no concordaran con el orden esperado para tres estímulos (“nunca”, “a veces”, “siempre”). De esta manera, se conservaron para el análisis los datos obtenidos de 44 participantes.

En la Tabla 1 se muestran los logaritmos de las medias geométricas obtenidas siguiendo el procedimiento descrito en el último párrafo de la sección de Método, pero separando los valores obtenidos según el referente usado como primer estímulo (“a veces” o “normalmente”). Las pruebas estadísticas (correlación y t de Student) reportadas en este artículo se aplicaron usando estos valores. Para verificar el nivel de convergencia entre los valores obtenidos con los dos posibles referentes, se correlacionaron los logaritmos base 10 de las medias geométricas. La correlación alta entre estos valores demuestra el alto grado de convergencia en las estimaciones cuantitativas dadas por los participantes ($r = .99, p < .001$).

Para corroborar la convergencia entre resultados obtenidos con los dos referentes, se compararon las puntuaciones dadas por los participantes a cada estímulo, con el fin de identificar si alguno de los estímulos presentó diferencias significativas. Como los valores dados por los participantes podían diferir al usar un referente distinto, se igualaron los valores (logaritmos) utilizando el procedimiento de igualación de extremos (ver Cañadas Osinski & Sánchez Bruno, 1998). Los valores así obtenidos fueron comparados con una t de Student para muestras relacionadas (ver Tabla 1), no se encontraron diferencias significativas en ninguno de los 27 estímulos.

Tabla 1. Media y desviación estándar de logaritmos de los 27 estímulos; comparación entre logaritmos obtenidos con dos referentes (“a veces”, “normalmente”). La última columna es la media geométrica conjunta que se utilizó para obtener valores intervalares.

Expresión	<i>a veces</i>		<i>normalmente</i>		Comparación		Media geométrica
	Media	DE	Media	DE	t	p	
Nada	-0.97	1.33	-0.90	1.42	-0.24	.81	0.12
Nunca	-0.90	1.37	-0.97	1.35	0.24	.81	0.12

Casi nunca	0.38	0.66	0.69	1.01	1.18	.24	3.44
Poco	0.61	0.76	0.71	0.57	-0.95	.35	4.58
Muy raramente	0.69	0.56	0.79	0.74	0.57	.57	5.52
Pocas veces	0.80	0.76	0.72	0.57	-0.90	.37	5.70
Raramente	0.72	0.74	0.80	0.66	-1.46	.15	5.78
Muy pocas veces	0.73	0.63	0.83	0.63	-1.16	.25	6.01
Rara vez	0.82	0.71	0.82	0.61	-0.26	.79	6.62
Raras veces	0.84	0.53	0.86	0.61	0.37	.72	7.10
Alguna vez	0.83	0.61	0.95	0.61	0.73	.47	7.71
De vez en cuando	1.07	0.48	1.14	0.61	1.04	.30	12.79
Algunas veces	1.09	0.49	1.15	0.58	-0.40	.69	13.20
Ocasionalmente	1.10	0.58	1.17	0.54	0.32	.75	13.63
A veces	1.11	0.54	1.17	0.55	0.74	.46	13.85
Ordinariamente	1.14	0.74	1.21	0.78	0.07	.95	14.97
A menudo	1.28	0.73	1.22	0.60	-1.04	.31	17.75
Normalmente	1.32	0.56	1.32	0.65	0.01	.995	20.88
Con frecuencia	1.30	0.54	1.39	0.61	0.36	.72	22.24
Muy a menudo	1.35	0.51	1.35	0.58	-0.41	.68	22.33
Generalmente	1.33	0.54	1.42	0.62	0.32	.75	23.94
La mayoría de las veces	1.42	0.74	1.34	0.59	-0.54	.59	23.97
Bastante	1.32	0.60	1.46	0.80	1.46	.15	24.53
Casi siempre	1.37	0.50	1.44	0.58	0.01	.996	25.51

Muchas veces	1.41	0.54	1.47	0.60	-0.68	.50	27.42
Muchísimas veces	1.46	0.76	1.47	0.64	-0.55	.58	29.13
Siempre	1.54	0.57	1.53	0.59	0.00	.9997	34.07

Una vez que se identificó la convergencia de los resultados con ambos referentes, se realizaron los cálculos de medias geométricas para cada cuantificador lingüístico utilizando los datos combinados obtenidos con los dos referentes (última columna de Tabla 1).

Las medias geométricas se tomaron como estimaciones cuantitativas del significado relativo de los 27 estímulos, por lo tanto, esto nos permitió seleccionar los estímulos más apropiados para la construcción de reactivos tipo Likert de diferente cantidad de puntos en la escala (entre 4 y 7). Para la selección de los estímulos, se tomaron los dos valores extremos (máximo y mínimo), y se calcularon los valores intermedios que se encontraran a intervalos iguales. A partir de estos valores calculados, se seleccionó el estímulo más cercano a esa puntuación, obteniendo las escalas que se presentan en la Tabla 2.

A manera de ejemplo, para la construcción de reactivos tipo Likert de 5 puntos, se recomienda utilizar las opciones: *nada/nunca*, *alguna vez*, *a menudo*, *casi siempre*, y *siempre*.

Tabla 2. Opciones de respuesta óptimas para diferente tamaño de escalas (4 a 7 puntos). La media geométrica real (columna izquierda) se compara con la media esperada para obtener intervalos iguales (4 columnas del lado derecho).

Expresión	Media geométrica	4 puntos	5 puntos	6 puntos	7 puntos
Nada/Nunca	0.12	0.12	0.12	0.12	0.12
Raramente	5.78				5.77
Raras veces	7.10			6.91	
Alguna vez	7.71		8.60		
De vez en cuando	12.79	11.43			11.43
Ocasionalmente	13.63			13.70	
A menudo	17.75		17.09		17.09

Normalmente	20.88			20.49	
Muy a menudo	22.33	22.75			22.75
Casi siempre	25.51		25.58		
Muchas veces	27.42			27.28	
Muchísimas veces	29.13				28.41
Siempre	34.07	34.07	34.07	34.07	34.07

4. CONCLUSIONES

El presente estudio demuestra la factibilidad y validez del uso de estimación de magnitudes para obtener puntuaciones intervalares de cuantificadores lingüísticos en México. Este método permite obtener recomendaciones de cuáles opciones de respuesta usar al construir escalas tipo Likert de frecuencia, por lo tanto, permite obtener un dato sobre la validez de las opciones utilizadas.

Aplicando el método de estimación de magnitudes, con una selección más específica de cuantificadores en una muestra más representativa de las regiones lingüísticas del país permitiría el obtener escalas tipo Likert a nivel de intervalos recomendables para México. El uso automatizado de la herramienta a través de la aplicación computarizada usado en el presente estudio facilita la recolección de datos y hace más factible una aplicación más amplia del método.

A partir de los resultados presentados sobre los valores intervalares de los cuantificadores de frecuencia, se espera que las recomendaciones obtenidas a partir del método de estimación de magnitudes sean utilizadas como referencia de mayor objetividad para la construcción de escalas que utilicen reactivos tipo Likert en la región.

REFERENCIAS

1. Boone, H. N. J., & Boone, D. a. (2012). Analyzing Likert Data. *Journal of Extension*, 50(2), 30. <http://doi.org/10.1111/j.1365-2929.2004.02012.x>
2. Cañadas Osinski, I., Prieto Marañón, P., San Luis Costas, C., & Caballero de Rodas, F. D. (1994). Estudio de cuantificadores lingüísticos de frecuencia y cantidad. *Acta Comportamental*, 2(2), 237–253.
3. Cañadas Osinski, I., & Sánchez Bruno, A. (1998). Categorías de respuesta en escalas tipo likert. *Psicothema*, 10(3), 623–631.
4. Cañadas Osinski, I., & Sánchez Bruno, A. (2000). Estudio de cuantificadores lingüísticos de cantidad aplicados a escalas psicológicas. *Revista de Psicología General Y Aplicada*, 53(3), 451–460.
5. Fontes, S., Fontes, A. I., & Merino, J. M. (1990). Comparación de dos técnicas psicofísicas sobre la

- función de Stevens. *Revista Latinoamericana de Psicología*, 22(3), 373–383.
<http://doi.org/10.4090/juee.2008.v2n2.033040>
6. Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating Ordinal Reliability for Likert-Type and Ordinal Item Response Data: A Conceptual, Empirical, and Practical Guide. *Practical Assessment, Research & Evaluation*, 17(3), 1–13. <http://doi.org/http://pareonline.net/getvn.asp?v=17&n=3>
 7. Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, 7(4), 396–403. <http://doi.org/10.9734/BJAST/2015/14975>
 8. Li, Q. (2013). A novel Likert scale based on fuzzy sets theory. *Expert Systems with Applications*, 40(5), 1609–1618. <http://doi.org/10.1016/j.eswa.2012.09.015>
 9. Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 1–55. <http://doi.org/2731047>
 10. Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314–324. <http://doi.org/10.3758/s13428-011-0168-7>
 11. Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), 625–632. <http://doi.org/10.1007/s10459-010-9222-y>
 12. Schriesheim, C. A., & Castro, S. L. (1996). Referent effects in the magnitude estimation scaling of frequency expressions for response anchor sets: an empirical investigation. *Educational and Psychological Measurement*, 56(4), 557–569.
 13. Schriesheim, C. A., & Gardiner, C. C. (1992). A comparative test of magnitude estimation and pair-comparison treatment of complete ranks for scaling a small number of equal-interval frequency response anchors. *Educational and Psychological Measurement*, 52, 867–877. <http://doi.org/10.1177/0146167299025006006>
 14. Schriesheim, C. A., & Schriesheim, J. F. (1978). The invariance of anchor points obtained by magnitude estimation and pair-comparison treatment of complete ranks scaling procedures: an empirical comparison and implications for validity of measurement. *Educational and Psychological Measurement*, 38, 977–983.
 15. Stevens, S. S. (1956). The direct estimation of sensory magnitudes-Loudness. *The American Journal of Psychology*, LXIX(1), 1–25.

UVCN-CL02

Validez factorial de escala de estrés (NNS) para enfermería en población mexicana

María de Lourdes Preciado Serrano, Araceli Ambriz Ramos, Gustavo Hernández Chávez

Universidad de Guadalajara - Centro Universitario de Ciencias de la Salud; malourdespre@gmail.com

RESUMEN

La teoría transaccional establece que el equilibrio entre las condiciones laborales y la respuesta emocional, favorece la satisfacción o la respuesta de estrés. Particularmente, la evaluación cognitiva que hace un trabajador, tanto de sus recursos personales en interacción con su historia vivida, le permite identificar los recursos que activará como respuesta ante un contexto determinado, considerándolo como estresor o no. Existen evidencias de que el profesional de enfermería, por su actividad en sí, es vulnerable a padecer estrés laboral, precisamente, por la exigencia del servicio especializado para el cuidado de la vida de otros, su constante contacto con el dolor y muerte de sus pacientes, entre otras cuestiones administrativas. Identificar la estructura factorial de la escala The Nursing Stress Scale (NSS) para valorar las situaciones de estrés laboral percibidas por el personal de enfermería de un hospital de