

¿Ítems Politómicos o Dicotómicos? Un estudio empírico con una escala unidimensional

Dominguez Lara, Sergio A. *

Facultad de Psicología y Trabajo Social, Universidad Inca Garcilaso de la Vega, Lima, Perú.

Artículo Metodológico

Resumen

Las escalas tipo Likert desde su aparición y difusión han tenido gran acogida y son consideradas como uno de los métodos más adecuados para construir tests psicológicos. Además, se asume que mientras más opciones tengan se obtendrá mayor información acerca del individuo y por tanto presentará mejores propiedades psicométricas. El objetivo de este trabajo fue evaluar la pertinencia del uso de diferentes formatos de respuestas dentro de una escala tipo Likert. Se realizaron análisis psicométricos en los diversos formatos propuestos en la escala (análisis de ítems, confiabilidad y estructura factorial). Se encontraron similitudes en cuanto a las características psicométricas de cuatro formatos propuestos. Se concluye que el formato establecido no afecta las propiedades psicométricas del instrumento como la confiabilidad y validez.

Palabras Claves:

Escala Tipo Likert; Formato de Respuesta; Confiabilidad; Validez.

Recibido el 1 de Agosto de 2013; Recibida la revisión el 10 de Octubre de 2013; Aceptado el 12 de Octubre de 2013.

Abstract

Polytomous or dichotomous items? An empirical study with an unidimensional scale: Likert scales from its apparition and diffusion have been well received and are considered one of the best methods to build psychological testing. Furthermore, it is assumed than if have more response options, you will have more information about the individual, and better psychometric properties. The aim of this study was to evaluate the relevance of the use of different formats for answers in a Likert-type scale. Were conducted psychometric analyzes in various formats proposed in the scale (item analysis, reliability and factor structure). Were found similarities in terms of the psychometric characteristics of four formats proposed. We conclude that both politomous and dichotomous formats provide important information as to the reliability and validity of the instrument.

Key Words:

Likert-Type Scale; Response Format; Reliability; Validity.

1. Introducción

Las escalas tipo Likert, o politómicas, desde su aparición y difusión en las primeras décadas del siglo pasado (Likert, 1932), han tenido gran acogida y sus postulados fueron tomados en cuenta para la construcción de muchos instrumentos de evaluación psicológica que ahora conocemos (Cortada de Kohan, 2004), ya que se asume que mientras más opciones tengan se obtendrá mayor información acerca del individuo (Comrey, 1988; Garner, 1960; Green & Rao, 1970), y por tanto se verán incrementadas tanto su confiabilidad (Boote, 1981; Oaster, 1989), su capacidad discriminativa, y su validez (Andrews, 1984; Nunnally & Bernstein, 1995).

En tal sentido, en las escalas tipo Likert se asume

que existe una equivalencia en cuanto la distancia que existe entre las opciones que plantea la escala y la manifestación del constructo en la persona. Es decir, si hay cuatro opciones de respuesta en la escala (*nunca, casi nunca, algunas veces y siempre*) para cualquier enunciado, la *distancia psicológica* entre *nunca* y *casi nunca* es la misma que entre *casi nunca* y *algunas veces*. Esto implica la existencia de *umbrales de respuesta* entre cada par de opciones planteadas en la escala (López, 2005), es decir, si tiene n opciones de respuesta, tendrá $n-1$ umbrales, de los cuales se espera que reflejen la *distancia psicológica* entre las opciones de respuesta planteadas, y se asume además la existencia de aquellos $n-1$ umbrales en las personas que

* Enviar correspondencia a: Dominguez Lara, S.A.
E-mail: sdominguezmpcs@gmail.com

responden dicho instrumento.

López (2005) plantea que esto es difícil de cumplir, ya que es probable que haya opciones de respuesta con frecuencias muy bajas (es decir, pocas personas que las hayas elegido como la opción más adecuada), o que nadie las haya contestado y, por lo tanto, umbrales que no existen. En ese caso, es más adecuado decidir por menos opciones que las planteadas inicialmente, y que reflejen mejor lo que se desea evaluar.

Por otro lado se encuentran los ítems en formato dicotómico, los cuales tienen un solo umbral, el cual asume “total ausencia” (No/Nunca/En desacuerdo/Fallo) o “total presencia” (Sí/Siempre/De acuerdo/Acierto) de aquella conducta que se refleja en el ítem que forma parte de la evaluación de algún constructo en particular.

Los estudios de simulación han jugado un rol fundamental en este tipo de temáticas. Bandalos y Enders (1996) encontraron que la confiabilidad fue mayor en las escalas que tienen entre 5 y 7 opciones de respuesta. Preston y Colman (2000) examinaron opciones de respuesta entre 2 y 11 puntos, y encontraron que la confiabilidad test-retest fue más baja en las escalas que tienen entre 2 y 4 opciones de respuesta, y fue mayor en las escalas entre 7 y 10 opciones de respuesta, y disminuía en escalas con más de 10 opciones de respuesta. Por su parte, Lozano, García-Cueto y Muñiz (2008) examinaron la confiabilidad y validez de escalas entre 2 y 9 opciones de respuesta en cuatro diferentes tamaños de muestra. Encontraron que el número óptimo de opciones de respuesta fue de 4 a 7, y que con menos de 4 opciones de respuesta la confiabilidad y validez (expresada mediante el porcentaje de varianza explicada) disminuye.

Por otro lado, se encuentran aquellos que postulan que el número de opciones no afecta la respuesta de los participantes (Chang, 1994; Schuts & Rucker, 1975). En este caso, se destaca el trabajo de López (2005), quien evaluó cinco formatos de opciones de respuesta (dos de 2 opciones, dos de 3 opciones, y uno de 4 opciones de respuesta) en un test de depresión aplicado a una muestra de hombres y mujeres norteamericanos. Concluye que emplear un formato politómico no mejora sistemáticamente la medida de depresión que un formato dicotómico, ya que se obtuvo la misma información psicométrica en torno a confiabilidad y validez.

Entonces, asumir de manera previa que los ítems politómicos de n opciones resultan más adecuados que

otras opciones sin considerar el tipo de instrumento ni el grupo al cual se está evaluando, es dejar de lado la posibilidad de una elección del formato de respuesta que posiblemente sea inadecuada, dado que los instrumentos varían según su marco teórico y sus objetivos de medición.

Cabe mencionar que si bien existe una clasificación de tests psicológicos subdivididos en *pruebas de habilidades*, *inventarios de personalidad* y tests que evalúan *constructos motivacionales* (Tornimbeni, Pérez, Olaz, & Fernández, 2004), resulta claro que para las *pruebas de habilidades* el formato dicotómico es el más adecuado, en tanto lo que se evalúa son las capacidades en torno a dos únicas opciones: acierto o error. Por el contrario, para las otras dos categorías, *inventarios de personalidad* y tests que evalúan *constructos motivacionales*, la elección del formato de repuesta debe hacerse con sumo cuidado, ya que es posible que puedan plantearse en formato politómico con n opciones, pero sea más adecuado un formato politómico con un número distinto de opciones, o uno dicotómico; o, por el contrario, plantearlo inicialmente de modo dicotómico, cuando en realidad no podría ser lo más adecuado.

Considerando lo revisado anteriormente, el objetivo de este trabajo fue analizar un instrumento estudiado por el equipo al cual pertenece el autor de este trabajo, la Escala de Autoeficacia Percibida Específica de Situaciones Académicas-EAPESA (Dominguez, Villegas, Yauri, Mattos, & Ramírez, 2012), mediante cinco formatos de puntuación a partir de las cuatro opciones de respuesta originales (el original, y cuatro adicionales), a fin de evaluar la pertinencia del uso tanto de ítems dicotómicos como politómicos, explorando las propiedades psicométricas (confiabilidad y validez) de la escala considerando los cinco formatos propuestos

2. Método

Según Montero y León (2007) se empleó un estudio instrumental, destinado al estudio de las propiedades psicométricas de un test.

2.1. Participantes

La muestra que participó en el estudio estuvo conformada por 235 estudiantes de psicología de una universidad privada, de segundo a séptimo ciclo, 60 varones y 175 mujeres, de edades comprendidas entre 16 y 37 años ($M = 20.86$). El muestreo utilizado fue de tipo intencional.

2.2. Instrumento

Escala de Autoeficacia Percibida Específica de Situaciones Académicas (EAPESA) de Palenzuela (1983) en versión de Dominguez et al. (2012). Consta de 9 ítems con cuatro opciones de respuesta (*Nunca, Algunas veces, Bastantes veces y Siempre*). Las propiedades psicométricas indican una consistencia interna adecuada (alfa de Cronbach de .89) y una estructura unidimensional, es decir, un solo factor, obtenido mediante el análisis paralelo en el trabajo original, que explica 55.26% de la varianza total del instrumento. El instrumento puede hallarse en el anexo.

2.3. Procedimiento

El inventario se administró como parte de una batería de evaluación psicológica dentro de la casa de estudios. Los responsables del estudio explicaron las instrucciones que figuran en el protocolo de aplicación de la prueba, y absolvieron las dudas que las personas evaluadas tuvieron.

Posteriormente, se elaboró una base de datos en MS Excel. A partir de dicha base, se construyeron cuatro formatos adicionales siguiendo las recomendaciones de López (2005). El formato 1 hace referencia a la distribución original de las opciones; el formato 2 parte del supuesto que las dos primeras opciones conforme una sola, siendo las otras dos independientes; el formato 3 plantea la unión de las últimas dos opciones, dejando las dos primeras de manera independiente.

Los dos últimos formatos hacen referencia a la posibilidad de una situación que indique dicotomía. En el formato 4 se considera que las dos primeras y las dos últimas opciones forman dos opciones independientes; y por último, el formato 5, la posibilidad de polaridad en la respuesta, en tanto la opción *Nunca* conforme una que contenga la *ausencia* de la conducta evaluada, y las otras tres opciones en una sola que índice que *presencia* de la conducta evaluada.

Tabla 1.

Formatos a partir de las opciones de respuesta originales del EAPESA

	A=Nunca	B=Algunas veces	C=Bastantes veces	D=Siempre
Formato 1	1	2	3	4
Formato 2	1	1	2	3
Formato 3	1	2	3	3
Formato 4	1	1	2	2
Formato 5	1	2	2	2

En cada uno de dichos formatos fue realizado un análisis de ítems (índices de homogeneidad), la

confiabilidad por consistencia interna, la estructura factorial y coeficientes de congruencia (García-Cueto, 1994). Los resultados de cada uno de tales procedimientos (medias aritméticas e índices de homogeneidad) fueron analizados de forma descriptiva, así como mediante el coeficiente de correlación de Pearson. Asimismo, se presentaron los coeficientes de consistencia interna y de congruencia.

El análisis de los datos se llevó a cabo utilizando el programa ViSta (Young, 2003) en lo que concierne al análisis de ítems y las correlaciones; FACTOR (Lorenzo-Seva & Ferrando, 2007) para la realización del análisis factorial; un módulo para el cálculo del alfa ordinal (Dominguez, 2012) y el cálculo del coeficiente de congruencia en MS Excel (Dominguez, en prensa).

3. Resultados

3.1. Estadísticos descriptivos

Se analizaron los estadísticos descriptivos de cada uno de los reactivos en los cinco formatos que se plantearon inicialmente. En los cuatro formatos iniciales se aprecian indicadores descriptivos dentro de rangos aceptables, más en el quinto formato se aprecian elevados indicadores de asimetría y curtosis, y varianzas cercanas a cero, lo que dificultó los análisis posteriores.

Por otro lado, las correlaciones efectuadas entre las medias de los cuatro formatos iniciales se encuentran por encima de .90 y son estadísticamente significativas. No obstante, con el quinto formato no se logró un panorama similar, obteniéndose indicadores desfavorables con todos los formatos restantes (del primero al cuarto).

Tabla 3.

Correlación entre las medias de los cinco formatos.

	F2	F3	F4	F5
F1	.998***	.937***	.958***	-.510
F2	-	.925***	.952***	-.559
F3		-	.993***	-.302
F4			-	-.414

Nota. F1: Formato Original; F2: Politémico 1; F3: Politémico 2; F4: Dicotómico 1; F5: Dicotómico 2. ***p < .001.

3.2. Análisis de ítems

En cuanto al análisis de los índices de homogeneidad del test, se examinó el grado de asociación entre los ítems que conforman la prueba y el test (Elosua, 2003), y se consideraron adecuados aquellos que obtuvieron una correlación ítem-test mayor de .20 (Likert, 1932). Como puede apreciarse en la tabla que contiene los resultados, los cuatro primeros formatos obtuvieron correlaciones ítem-total por

encima del mínimo requerido. El quinto formato no admitió dicho análisis.

Meyer-Olkin (KMO) y el test de Esfericidad de Bartlett. Después de ello, se llevó a cabo un análisis factorial por *mínimos cuadrados no ponderados*, usando matrices tetracóricas y policóricas (no las matrices Pearson comúnmente usadas, dado que no corresponden para variable de tipo ordinal y nominal). No se efectuó rotación alguna en vista que se extraerá un solo factor.

De acuerdo con los resultados, los cuatro formatos estudiados presentan indicadores adecuados previos al análisis factorial que justifican su procedimiento (KMO adecuado y Test de Esfericidad de Bartlett significativo), aunque cabe destacar que el indicador más favorable lo presenta el Formato 2. Asimismo, las cargas factoriales presentan magnitudes superiores a .40 (Glutting, 2002; Zwick & Velice, 1986) y la varianza explicada en cada formato estudiado supera el 50%, lo cual indica unidimensionalidad en todos ellos (Carmines & Zeller, 1979). La estructura factorial del Formato 2 fue la que presentó cargas factoriales de mayor magnitud y mayor varianza explicada (64.05%).

La confiabilidad por medio de sus consistencia interna se estimó para cada uno de los formatos establecidos mediante el alfa ordinal (Dominguez, 2012; Elosua & Zumbo, 2008) y el KR20, en el caso del Formato 4, los cuales están entre .80 y .90, que se consideran adecuados (Campo-Arias & Oviedo, 2008; Hogan, 2004), destacándose el Formato 2 que presenta el indicador de mayor magnitud.

Tabla 5.

Correlación entre los índices de homogeneidad de los formatos de respuesta.

	F2	F3	F4
F1	.994***	.985***	.968***
F2	-	.973***	.975***
F3		-	.983***

Nota. F1: Formato Original; F2: Politémico 1; F3: Politémico 2; F4: Dicotómico 1; F5: Dicotómico 2. ***p < .001.

3.4. Comparación de estructuras factoriales

Se aprecian coeficientes de congruencia cercanos a 1.00 y estadísticamente significativos (García-Cueto, 1994).

4. Discusión

El objetivo de este trabajo fue evaluar la pertinencia del uso de diferentes formatos de respuestas dentro de una escala tipo Likert, a fin de contrastar los planteamientos antecedentes tanto a favor como en contra de un incremento (o decremento) del número de opciones para mejorar la confiabilidad y validez del instrumento.

El primer punto estudiado estuvo enfocado en los

Tabla 2. Estadísticos descriptivos de los cinco formatos de respuesta

	F1			F2			F3						
	M	DE	Asm	Crt	M	DE	Asm	Crt	M	DE	Asm	Crt	M
Ítem 1	2.991	.815	-.269	-.800	2.017	.766	-.029	-1.289	2.689	.515	-1.373	.901	1.715
Ítem 2	2.881	.763	-.084	-.673	1.902	.722	.150	-1.072	2.664	.516	-1.157	.255	1.685
Ítem 3	2.847	.800	.083	-1.032	1.864	.771	.239	-1.277	2.609	.522	-0.805	-.593	1.626
Ítem 4	2.881	.752	.079	-.957	1.889	.736	.177	-1.132	2.660	.492	-.891	-.681	1.668
Ítem 5	2.860	.856	.029	-1.166	1.885	.814	.214	-1.455	2.579	.543	-.797	-.469	1.604
Ítem 6	2.983	.760	-.088	-.967	1.991	.744	.014	-1.184	2.711	.472	-1.175	-.014	1.719
Ítem 7	3.247	.792	-.831	.103	2.277	.723	-.474	-.978	2.809	.463	-2.428	5.218	1.838
Ítem 8	2.860	.778	.142	-1.104	1.868	.764	.228	-1.248	2.626	.501	-.724	-.986	1.634
Ítem 9	2.966	.804	-.235	-.754	1.991	.755	.014	-1.238	2.685	.517	-1.347	.823	1.711

Nota. M = media; DE = desviación estándar; Asm = asimetría; Crt = curtosis; F1: Formato Original; F2: Politémico 1; F3: Politémico 2; F4: Dicotómico 1; F5: Dicotómico 2.

3.3. Análisis Factorial

Para cada uno de los formatos propuestos, se exploró si los ítems poseían la suficiente correlación para realizar el análisis factorial. Esto se realizó mediante el Índice de Adecuación Muestral de Kaiser-

estadísticos descriptivos de cada uno de los cinco formatos propuestos inicialmente. Al respecto, los Formatos 3 y 5, de tres y dos opciones respectivamente, fueron los que obtuvieron los indicadores de asimetría y curtosis más pronunciados de todo el análisis. Esto se debe a que la opción de respuesta *Nunca*, tuvo una frecuencia baja en todos los formatos propuestos, pero en los Formatos 3 y 5 conformó una sola opción de respuesta. Si bien en el Formato 1 también estuvo

aislada, la existencia de la opción *Casi Nunca*, atenuó su presencia, mas en el Formato 2 y 4 estuvo con la opción *Casi Nunca* en una sola opción. No obstante, el aislamiento en una sola opción de respuesta en el Formato 5, llevó a tal formato a la eliminación posterior debido a que su media aritmética difería en milésimas del puntaje máximo posible (lo cual hacía su varianza casi cero) y su asimetría y curtosis fueron exorbitantes.

Tabla 4.

Análisis de ítems, Análisis Factorial y Confiabilidad en los formatos de respuesta.

	Índices de Homogeneidad				Carga Factorial			
	F1	F2	F3	F4	F1	F2	F3	F4
Ítem 1	.653	.664	.546	.562	.669	.742	.717	.746
Ítem 2	.659	.667	.542	.558	.656	.741	.706	.727
Ítem 3	.727	.726	.618	.603	.710	.808	.754	.769
Ítem 4	.720	.716	.640	.635	.700	.790	.788	.814
Ítem 5	.694	.713	.581	.620	.689	.799	.726	.801
Ítem 6	.776	.779	.673	.678	.846	.856	.829	.866
Ítem 7	.571	.595	.415	.448	.685	.672	.665	.643
Ítem 8	.626	.637	.499	.514	.625	.712	.663	.675
Ítem 9	.727	.737	.650	.662	.770	.815	.806	.846
KMO	-	-	-	-	.860	.940	.900	.870
Test de Bartlett (p-valor)	-	-	-	-	<.001	<.001	<.001	<.001
% Varianza explicada	-	-	-	-	55.53%	64.05%	59.84%	63.41%
Alfa Ordinal y KR20	-	-	-	-	.899	.929	.915	.927

Nota. F1: Formato Original; F2: Politémico 1; F3: Politémico 2; F4: Dicotómico 1; F5: Dicotómico 2. ***p < .001.

Tabla 6.

Coeficientes de congruencia entre los formatos de respuesta.

	F2	F3	F4
F1	.996***	.995***	.996***
F2	-	.983***	.989***
F3		-	.998***

Nota. F1: Formato Original; F2: Politémico 1; F3: Politémico 2; F4: Dicotómico 1; F5: Dicotómico 2. ***p < .001.

En segundo lugar, una vez descartado el quinto formato, con los cuatro formatos iniciales se procedió a realizar un análisis de ítems, empleando la correlación ítem-test en los Formatos 1, 2 y 3, y la correlación biserial-puntal para Formato 4 (ítems dicotómicos). El quinto formato no admitió dicho análisis por los indicadores de dispersión que poseía. Los indicadores obtenidos en cada uno de los formatos son aceptables, ya que están por encima de .20, aunque en términos generales el Formato 2 presenta indicadores más sólidos, y el Formato 3, los indicadores más bajos. Asimismo, se apreciaron correlaciones significativas entre los índices de homogeneidad de los cuatro formatos, lo cual indica que al margen del formato usado, cada ítem evalúa aspectos similares que los

demás. Esto concuerda con los hallazgos de López (2005), quien también encontró correlaciones elevadas entre los diversos formatos de respuesta con relación al índice de homogeneidad.

Además de ello, se estudiaron las estructuras factoriales de los cuatro formatos propuestos. Cada una de ellas refleja la unidimensionalidad del constructo, evidenciando entonces la validez del instrumento al margen del número de opciones de respuesta planteado inicialmente. En tal sentido, son los Formatos 2 y 4, los que presentan los indicadores más elevados de varianza explicada, siendo el Formato 1 (original) el que presenta el porcentaje más bajo de todos (55.53%). Esto difiere, junto con el estudio de López (2005), de aquellos estudios que plantean que mientras menos opciones de respuesta tengan los ítems disminuirá su validez.

Siguiendo con la estructura factorial, se realizó un análisis de congruencia entre los factores hallados en cada formato. Se compararon las estructuras factoriales a fin de explorar si los cuatro formatos están midiendo la misma dimensión o no, ya que es lo que se espera, y si dicha congruencia es estadísticamente significativa o no. En tal sentido, se hallaron coeficientes de

congruencia cercanos a 1.00 y estadísticamente significativos (García-Cueto, 1994), lo cual da cuenta de la semejanza de las estructuras factoriales al margen del número de opciones de respuesta con el que se cuente. Esto apoya lo planteado en el párrafo anterior.

Por último, con relación a la confiabilidad, su calidad no se ve alterada si varían los formatos de puntuación, aunque cabe resaltar que los Formatos 2 y 4, son los que tienen indicadores más aceptables.

Entonces, a partir de los hallazgos presentados, se puede concluir que tanto los formatos politómicos como dicotómicos no afectan las propiedades psicométricas del instrumento como la confiabilidad y validez del instrumento.

Entre las limitaciones se destaca que en este estudio se analiza un instrumento en particular, en el cual se observa una frecuencia no significativa en una de sus opciones de respuesta (*Nunca*), por lo cual los hallazgos deben ser tomados con precaución. Cabe mencionar que López (2005) utilizó un instrumento con sesgo negativo, y obtuvo resultados similares a los reportados aquí.

Por otro lado, con la eliminación del quinto formato no se plantea excluir el formato de respuesta dicotómico, ya que como se comprobó, uno de los formatos dicotómicos (el Formato 4) tiene indicadores aceptables; lo que se rescata es que en el instrumento que evalúa este constructo (autoeficacia), es poco probable que la opción *Nunca* sea respondida de forma frecuente, al contrario del instrumento que evaluó López (2005), en el cual se evalúan aspectos depresivos. Esto último indica la necesidad de considerar a fondo las características del test psicológico al cual se pretende asignar opciones de respuesta.

Además, el instrumento estudiado no es representativo de otros instrumentos o pruebas, por lo cual los hallazgos no pueden ser generalizados a otros tests psicológicos. Asimismo, no se estudiaron fuentes externas de validez. Se recomienda, para futuras investigaciones, estudiar tests psicológicos multidimensionales para análisis de esta naturaleza, ya que si bien se está encontrando evidencia que da sustento al uso de diversas opciones de respuesta, esta se refiere a un instrumento unidimensional. Del mismo modo, se recomienda implementar diseños inter-grupos, en los cuales se evalúen los diferentes formatos de respuesta en distintas muestras.

Referencias

- Andrews, F. (1984). Construct validity and error components of survey measures: A structural modeling approach. *Public Opinion Quarterly*, 48, 409-442.
- Bandalos, D., & Enders, C. (1996). The effects of non-normality and number of response categories on reliability. *Applied Measurement in Education*, 9, 151-160.
- Boote, A. (1981). Reliability testing of psychographic scales: Five-point or seven-point? Anchored or labeled? *Journal of Advertising Research*, 21, 53-60.
- Campo-Arias, A., & Oviedo, H. (2008). Propiedades psicométricas de una escala: la consistencia interna. *Revista de Salud Pública*, 10(5), 831-839.
- Carmines, E., & Zeller, R. (1979). *Reliability and validity assessment*. London: Sage.
- Chang, L. (1994). A psychometric evaluation of four point and six point Likert-type scales in relation to reliability and validity. *Applied Psychological Measurement*, 18, 205-215.
- Comrey, A. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, 56, 754-761.
- Cortada de Kohan, N. (2004). *Teoría y métodos para la construcción de escalas de actitudes*. Buenos Aires: Lugar Editorial.
- Dominguez, S. (2012). Propuesta para el cálculo del Alfa Ordinal y Theta de Armor. *Revista de Investigación en Psicología*, 15(1), 213-217.
- Dominguez, S. (en prensa). Coeficiente de congruencia: propuesta de cálculo. *Pensando Psicología. En prensa*.
- Dominguez, S., Villegas, G., Yauri, C., Mattos E., & Ramírez F. (2012). Propiedades psicométricas de una escala de autoeficacia para situaciones académicas en estudiantes universitarios peruanos. *Revista de Psicología-Universidad Católica San Pablo*, 2(1), 27-39.
- Elosua, P. (2003). Sobre la validez de los tests. *Psicothema*, 15(2), 315-321.
- Elosua, P., & Zumbo, B. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema*, 20(4), 896-901.
- Garner, W. (1960). Rating scales, discriminability and information transmission. *Psychological Review*, 67, 343-352.
- García-Cueto, E. (1994). Coeficiente de congruencia. *Psicothema*, 6(3), 465-468.
- Glutting, J. (2002). Some psychometric properties of a system to measure ADHD. *Measurement and Evaluation in Counseling and Development*, 34, 194-209.
- Green, P., & Rao, V. (1970). Rating scale and information recovery: How many scales and response category to use? *Journal of Marketing*, 34, 33-39.
- Hogan, T. (2004). *Pruebas psicológicas. Una introducción práctica*. México D.F.: Manual Moderno
- Likert, R. (1932). A Technique for the Measurement of

- Attitudes. *Archives of Psychology*, 140, 1-55.
- López, J. (2005). Items politómicos vs. dicotómicos: un estudio metodológico. *Anales de Psicología*, 21(2), 339-344.
- Lorenzo-Seva, U., & Ferrando, P. (2007). *FACTOR: A computer program to fit the exploratory factor analysis model*. University Rovira i Virgili.
- Lozano, L.M., Garcia-Cueto, E., & Muniz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *European Journal of Research Methods for the Behavioral and Social Sciences*, 4, 73-79.
- Montero, O., & León, I. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862.
- Nunnally, J., & Bernstein, I. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Oaster, T. (1989). Number of alternatives per choice point and stability of Likert-type scales. *Perceptual and Motor Skills*, 68, 549-550.
- Palenzuela, D. (1983). Construcción y validación de una escala de autoeficacia percibida específica de situaciones académicas. *Análisis y Modificación de Conducta*, 9(21), 185-219.
- Preston, C., & Colman, A. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104, 1-15.
- Schuts, H., & Rucker, M. (1975). A comparison of variables configuration across scale lengths: an empirical study. *Educational and Psychological Measurement*, 35, 319-324.
- Tornimbeni, S., Pérez, E., Olaz, F., & Fernández, A. (2004). *Introducción a los tests psicológicos*. Córdoba: Brujas.
- Young, F. (2003). ViSta "The Visual Statistics System". [programa informático]. Recuperado de: URL: <http://forrest.psych.unc.edu/research/index.html>, el 16.11.11, 2:00 p.m.
- Zwick, W., & Velicer, W. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, 432-442.

ANEXO

Escala de Autoeficacia Académica

A continuación encontrarás una serie de enunciados que hacen referencia a tu modo de pensar. Lee cada frase y contesta marcando con una X de acuerdo con la siguiente escala de valoración:

N = Nunca AV = Algunas veces B = Bastantes veces S = Siempre

¡POR FAVOR, NO DEJES NINGUNA PREGUNTA SIN RESPONDER!

	N	AV	B	S
1. Me considero lo suficientemente capacitado para enfrentarme con éxito a cualquier tarea académica				
2. Pienso que tengo capacidad para comprender bien y con rapidez una materia				
3. Me siento con confianza para abordar situaciones que ponen a prueba mi capacidad académica				
4. Tengo la convicción de que puedo obtener buenos resultados en los exámenes				
5. No me importa que los profesores sean exigentes y duros, pues confío en mi propia capacidad académica				
6. Creo que soy una persona capacitada y competente en mi vida académica				
7. Si me lo propongo, creo que tengo la suficiente capacidad para obtener un buen record académico				
8. Pienso que puedo pasar los cursos con bastante facilidad, e incluso, sacar buenas notas				
9. Creo que estoy preparado/a y capacitado/a para conseguir muchos éxitos académicos				