

XVI Congreso Internacional de Investigación y Práctica Profesional en Psicología. XXXI Jornadas de Investigación. XX Encuentro de Investigadores en Psicología del MERCOSUR. VI Encuentro de Investigación de Terapia Ocupacional. VI Encuentro de Musicoterapia. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires, 2024.

Faceta orden del modelo de los cinco factores. Simulaciones de su medición adaptativa.

Abal, Facundo Juan Pablo, Jares, Agustín Gabriel, Galibert, María Silvia y Lozzia, Gabriela.

Cita:

Abal, Facundo Juan Pablo, Jares, Agustín Gabriel, Galibert, María Silvia y Lozzia, Gabriela (2024). *Faceta orden del modelo de los cinco factores. Simulaciones de su medición adaptativa. XVI Congreso Internacional de Investigación y Práctica Profesional en Psicología. XXXI Jornadas de Investigación. XX Encuentro de Investigadores en Psicología del MERCOSUR. VI Encuentro de Investigación de Terapia Ocupacional. VI Encuentro de Musicoterapia. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires.*

Dirección estable: <https://www.aacademica.org/000-048/76>

ARK: <https://n2t.net/ark:/13683/evo3/FkY>

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

FACETA ORDEN DEL MODELO DE LOS CINCO FACTORES. SIMULACIONES DE SU MEDICIÓN ADAPTATIVA

Abal, Facundo Juan Pablo; Jares, Agustín Gabriel; Galibert, María Silvia; Lozzia, Gabriela
Universidad de Buenos Aires. Facultad de Psicología. Buenos Aires, Argentina.

RESUMEN

La faceta Orden del dominio Concienciación del Modelo de los Cinco Factores de la personalidad normal describe la tendencia de las personas a mantener su entorno personal organizado y ordenado, así como a establecer y seguir un horario de tareas. En investigaciones anteriores, se construyó una escala de 9 ítems para su medición y se calibró con el Modelo de Respuesta Gradual de la Teoría de Respuesta al Ítem. En el presente estudio, se realizan simulaciones de pruebas adaptativas para obtener una medida eficiente de la faceta. Participaron 441 sujetos de población general (51% mujeres) que respondieron a todos los ítems de la escala de Orden. Se utilizó el programa Firestar para programar el algoritmo que simula la administración adaptativa. Se compararon dos criterios para detener la administración: a) tras administrar 4 ítems y b) tras alcanzar un error de estimación del rasgo de 0,35 o responder a 5 ítems. Las correlaciones entre las estimaciones de rasgo obtenidas con la escala completa y las versiones adaptadas fueron altas ($r > .89$). Los resultados favorecieron ligeramente el modo de administración adaptado que planteaba un criterio de parada mixto.

Palabras clave

Faceta Orden - Modelo Cinco Factores - Test Adaptativo

ABSTRACT

ORDER FACET OF THE FIVE-FACTOR MODEL.
SIMULATIONS OF ITS ADAPTIVE MEASUREMENT

The Order facet of the Conscientiousness domain of the Five Factor Model of normal personality describes people's tendency to keep their personal environment organized and neat, as well as to set up and follow a task schedule. In previous research, a 9-item scale was constructed for its measurement and calibrated with the Graded Response Model of Item Response Theory. In the present study, adaptive testing simulations are performed to obtain an efficient measure of the facet. A total of 441 general population subjects (51% female) participated and responded to all items of the Order scale. The Firestar program was used to program the algorithm that simulates adaptive administration. Two criteria for stopping the administration were compared: a) after 4 items were administered and b) after an error of trait estimate of 0.35 was reached or 5 items were answered. Correlations between trait estimates obtained with the full scale and the adapted versions were high ($r > .89$). The results slightly

avored the adaptive mode of administration that posed a mixed stopping criterion.

Keywords

Order Facet - Five Factor Model - Adaptive Testing

La faceta Orden adquirió relevancia conceptual en el marco del Modelo de los Cinco Factores de la Personalidad - *Five Factor Model*- (FFM, Goldberg, 1993; McCrae y Costa, 2010) como una de las subdimensiones del dominio Responsabilidad. Se trata de la única faceta de este dominio que aparece sistemáticamente medida en las diversas operacionalizaciones del modelo y que, incluso, su contenido muestra una preponderancia destacada en los cuestionarios que miden la Responsabilidad de manera unidimensional (Sanchez Gonzalez y Galibert, 2022). Este rasgo describe la tendencia de las personas a conservar el entorno personal organizado y aseado, así como también a armar y seguir una planificación de las tareas. Los sujetos con niveles elevados de Orden suelen preocuparse excesivamente por la limpieza y el orden. Por el contrario, en los niveles bajos se encuentran las personas descuidadas y poco sistemáticas al organizar su vida (Green et al., 2015; Costa et al., 1991).

Recientemente se comenzó con la construcción de un banco de ítems de Responsabilidad en el que se incluyó un conjunto de elementos para medir Orden (Sánchez González et al., 2023). La escala consta de nueve ítems que fueron calibrados con el Modelo de Respuesta Graduada de la Teoría de Respuesta al Ítem. Los ítems muestran una adecuada capacidad discriminativa y parámetros de umbral que permiten localizar a los elementos en un nivel amplio del rasgo. A pesar de que la extensión de esta escala es adecuada y se ajusta a los formatos de las pruebas tradicionales como NEO-PI-3 (McCrae y Costa, 2010) o NEO-IPIP (Goldberg et al., 2006), en los últimos años se observa una tendencia hacia la construcción de test más y más breves (e.g. Soto & John, 2017) que ayuden a condiciones de evaluación más extremas como son los estudios a gran escala.

Dado que la reducción de los instrumentos erosiona indefectiblemente las garantías de calidad psicométricas, los Test Adaptativos Informatizados (TAIs) aparecen como una variante a considerar porque que permiten una medición eficiente de los rasgos sin comprometer la confiabilidad o validez de la medida (Reise y Moore, 2023). Un TAI se basa en Banco de Ítems que se usa de insumo para elegir el subconjunto de elementos más

eficientes para evaluar a un sujeto. Esta selección se programa a través de un algoritmo que define una modalidad de arranque, criterio para la elección progresiva de ítems y una modalidad de interrupción (Weiss y Sahin, 2024).

La evaluación de rasgos de la personalidad a partir de mediciones adaptativas registra un creciente interés a nivel mundial durante la última década (Reise y Revicki, 2015) aunque en Argentina se encuentra en un período incipiente tanto desde aspectos teóricos como aplicados (e.g. Abal et al., 2019; Abal et al., 2020). En este trabajo se propone ensayar la implementación de mediciones adaptativas de la escala que mide la faceta Orden con el objetivo de analizar la variante óptima para interrumpir la administración.

Método

Participantes

Colaboraron 441 personas adultas que vivían en el área metropolitana de Buenos Aires. Se llevó adelante un muestreo no probabilístico por accesibilidad. El 51% de los participantes consignó pertenecer al sexo femenino. La edad osciló entre 18 y 78 años con una edad promedio de 36.1 (DE= 15.7). El 56.3% de las personas alcanzó un nivel secundario completo como máximo y un 47.8% de los individuos reportó estar soltero.

Instrumentos

Ítems de la faceta Orden (Sánchez González et al., 2023). Consta de 9 ítems que evalúan el rasgo según la definición de Orden de McCrae y Costa (2010) y presentan un formato de respuesta de cuatro opciones tipo Likert (*En desacuerdo, Ligeramente en desacuerdo, Ligeramente de acuerdo, De acuerdo*). La escala presenta evidencias de validez que confirman una estructura interna unidimensional y ajustan al Modelo de Respuesta Graduada de Samejima (2016). En lo que respecta a la confiabilidad, registra índices de consistencia interna óptimos (Alfa de Cronbach = .83; Omega Ordinal = .91; Alfa Ordinal = .87).

Procedimiento

Los datos se recogieron a partir de evaluaciones individuales y sin tiempo límite. Previa participación se requirió la firma de un consentimiento informado el cual explicitaba los objetivos de la investigación, el carácter voluntario de la colaboración y la posibilidad de abandonar la actividad en cualquier momento si así lo deseaba. Asimismo, se brindaron garantías de anonimato y confidencialidad de sus respuestas. Si bien los sujetos respondieron el banco de Responsabilidad en su totalidad, para esta presentación solo se usaron las respuestas a la faceta Orden.

Análisis de datos

Considerando que en estudios previos ya se contaba con los ítems de Orden ya calibrados (Sánchez González et al., 2023), se procedió a realizar las simulaciones de administraciones adap-

tativas con los datos recolectados en esta nueva etapa con el software Firestar-D (Choi, 2018). El programa emplea la matriz de datos con las respuestas de los sujetos a todos los ítems y utiliza un algoritmo para procesar progresivamente aquellas respuestas que se le presentarían al evaluado si estuviera respondiendo un TAI.

El algoritmo adaptativo se especificó con un criterio de arranque en el que se considera la media del rasgo como la estimación inicial más probable. Luego de recuperar la respuesta a cada ítem, la estimación provisional de θ se efectuó con el método bayesiano de estimación esperada a posteriori (EAP) utilizando a la normal estándar como distribución a priori. En la selección progresiva de los ítems se utilizó el método de máxima información de Fisher. Finalmente, se obtuvo el error de estimación del nivel de cada evaluado en un formato de administración tradicional de 9 ítems y se analizó la precisión de la medida comparándolo con dos variantes adaptativas que se basaron en diferentes criterios de interrupción: a) test de longitud fija (en la que se administran 4 de los 9 ítems) y b) test de con criterio mixto (en la que se administran ítems hasta alcanzar un error mínimo de estimación del rasgo igual o menor a 0.35 - equivalente a una confiabilidad clásica de .88 aproximadamente - o, en su defecto, se interrumpe una vez respondido cinco ítems). Para analizar si las modalidades de parada brindan una estimación del rasgo similar a la que se obtendría a partir de la respuesta a los 9 ítems, se correlacionaron los niveles estimados bajo cada condición con los estimados al contestar el instrumento y con los puntajes brutos calculados tradicionalmente en el marco de la teoría clásica de tests. Adicionalmente, se analizó el error de estimación medio para la versión adaptativa de longitud fija; mientras que para la versión de longitud variable se obtuvo como indicador de su eficiencia la cantidad promedio de ítems que se requirió para alcanzar el nivel de error prefijado.

Resultados

Administración convencional

La administración de los 9 ítems que componen la escala arrojó una mediana en la estimación del rasgo de 0.39, oscilando entre -2.10 y 2.07. Para los casos aquí analizados el error de estimación registró una mediana de 0.28 por lo que es factible concluir que, en términos globales, la evaluación con todos los ítems presentó una confiabilidad clásica de aproximadamente .92.

El error de estimación del valor del rasgo más bajo de esta muestra fue de 0.22. Este dato indica que la base mínima de error que se podría alcanzar al medir el nivel de Orden de una persona, ya que la medición más precisa del rasgo se logra al administrar todos los ítems que componen la escala. En contraposición, el error máximo presentó un valor de 0.44 (equivalente a una confiabilidad clásica de .81). Con este resultado se observa que incluso los niveles del rasgo estimados con más error presentan una precisión elevada según los criterios clásicos.

Administraciones adaptativas

Al administrar de manera adaptativa 4 de los 9 ítems de la escala (versión de longitud fija) se obtuvo una mediana en el nivel estimado de rasgo de 0.35 con un mínimo de -1.84 y máximo de 1.78. La mediana del error de estimación bajo esta condición fue 0.34 (mínimo = 0.26 y máximo = 0.53), lo que resulta equivalente a una confiabilidad clásica aproximada de .88. Esto indica que, al administrar únicamente 4 ítems seleccionados de manera adaptativa se verifica un error levemente superior al obtenido con los 9 elementos (0.28 *versus* 0.34). Además, los valores ρ estimados con los 4 ítems correlacionaron de manera intensa y con la estimación con 9 ítems ($r = .88, p < .001$) y con el puntaje bruto calculado con teoría clásica ($r = .88, p < .001$). Los hallazgos con relación al TAI con un paradigma mixto muestran una mediana del nivel estimado de rasgo de 0.37 (mínimo = -1.87 y máximo = 1.97). Los errores de estimación oscilaron entre 0.28 y 0.50 con una mediana 0.34. Como es posible observar, el error de estimación se diferencia del obtenido con los 9 ítems en la misma magnitud que la registrada para el test de longitud fija (0.28 *versus* 0.34). No obstante, los puntajes estimados bajo esta variante se asociaron de manera ligeramente más intensa la versión completa ($r = .90, p < .001$) que las obtenidas con el test de longitud fija pero se mantuvo la intensidad con los puntajes brutos ($r = .88, p < .001$).

En esta variante adaptativa que propone un criterio mixto se requirió que los sujetos de la muestra respondan, en promedio, 4.21 ítems ($DE = 0.85$ ítems), con un mínimo de 2 y un máximo fijado de 5 ítems. El 69% de los participantes alcanzaron el error de 0.35 antes de presentarles el quinto ítem, lo que significa que un 31% no alcanzó el error prefijado y se interrumpió la administración por alcanzar el tope de cantidad de ítems.

Discusión

La evaluación de la personalidad atraviesa un período en el que se busca aumentar la eficiencia de las medidas ahorrando al máximo el tiempo de administración y conservando las propiedades psicométricas de las pruebas. En esta línea, los TAI se ofrecen como un recurso metodológico que posibilita disponer del tiempo no empleado para evitar la fatiga de los evaluados o para incluir más variables en un estudio a gran escala.

Las variantes de evaluación adaptativa simuladas para la faceta Orden han revelado que es posible obtener estimaciones de los niveles del rasgo con un grado de precisión óptimo aun cuando se administran apenas 4 o 5 ítems de la escala. Aunque ambas modalidades presentaron resultados similares, el criterio mixto para interrumpir el instrumento dejó en evidencia que brindar la posibilidad de administrar un ítem más favorece a un grupo de personas (aproximadamente 30%), para las cuales se reduce el error de estimación. Esto propicia que la correlación con el puntaje estimado a partir de la escala completa alcance una mayor intensidad.

Conviene señalar que la reducción en la cantidad de ítems administrados no supone un perjuicio sobre la representatividad del muestreo de contenido. En efecto, la correlación intensa entre el rasgo estimado con el TAI y la escala completa así lo demuestra. El TAI permite flexibilizar la evaluación para que cada sujeto sea medido con el conjunto de ítems que más información aporta a su nivel de rasgo. La infra o supra representación de un aspecto específico constitutivo de la faceta Orden originada por la medición adaptativa no tuvo consecuencias en la estimación. Incluso, la incidencia de esta variación podría ser menor que la ocasionada por un test breve con formato fijo, en el que todos los sujetos responden exactamente a los mismos contenidos. Futuros estudios abordarán otras potenciales modalidades de administración adaptativa. La inclusión de nuevos ítems al banco podría reducir incluso aun más la extensión de la prueba y reducir el error en la estimación del rasgo. Los hallazgos obtenidos son alentadores y muestran una metodología relevante que podría extenderse a las demás facetas del dominio.

BIBLIOGRAFÍA

- Abal, F. J. P., Auné, S. E., y Attorresi, H. F. (2019). Construcción de un banco de ítems de facetas de neuroticismo para el desarrollo de un test adaptativo. *Psicodebate*, 19(1), 31-50. <https://doi.org/10.18682/pd.v19i1.854>
- Abal, F. J. P., Sánchez González, J. F., Lozzia, G. S., y Attorresi, H. F. (2020). Escala de Desesperanza de Beck (BHS): ventajas de una administración adaptativa. *Revista Iberoamericana de psicología*, 14(1), 71-82. <https://doi.org/10.33881/2027-1786.rip.14106>
- Choi, S. W. (2018). Firestar. En W. J. van der Linden (Ed.). *Handbook of Item Response Theory, Volume 3: Applications* (pp. 547-556). Boca Raton: Chapman y Hall/CRC.
- Costa Jr, P. T., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences*, 12(9), 887-898. [https://doi.org/10.1016/0191-8869\(91\)90177-d](https://doi.org/10.1016/0191-8869(91)90177-d).
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American psychologist*, 48(1), 26.
- Green, J.A., O'Connor, D.B., Gartland, N., & Roberts, B.W. (2015) The[1] Chernyshenko Conscientiousness Scales. *Assessment*, 23(3), 374-385. <https://doi.org/10.1177/1073191115580639>.
- McCrae, R. R., & Costa, P. T., (2010). *NEO Inventories Professional Manual*. Odessa. Psychological Assessment Resources.
- Reise, S. P. y Revicki, D. A. (2015). *Handbook of Item Response Theory Modeling Applications to Typical Performance Assessment*. Nueva York: Routledge
- Reise, S. P., y Moore, T. M. (2023). Item response theory. En H. Cooper, M. N. Coutanche, L. M. McMullen, A. T. Panter, D. Rindskopf, y K. J. Sher (Eds.), *APA handbook of research methods in psychology: Foundations, planning, measures, and psychometrics* (2nd ed., pp. 809-835). American Psychological Association. <https://doi.org/10.1037/0000318-037>



- Samejima, F. (2016). Graded response models. En W. J. van der Linden (Ed.), *Handbook of Item Response Theory, Volume One*(pp. 123-136). Boca Raton: Chapman and Hall/CRC.
- Sanchez Gonzalez, J. F. & Galibert, M. S. (2022). Diferencias Teóricas en la Operacionalización de las Facetas del Dominio Responsabilidad del Modelo de los Cinco Grandes. *Memorias XIV Congreso Internacional de Investigación y Práctica Profesional en Psicología. Facultad de Psicología Universidad de Buenos Aires.*
- Sanchez Gonzalez, J. F., Auné, S., Attorresi, H. & Lozzia, G. S. (2023). Medición de la faceta Orden del Modelo de los Cinco Grandes usando Teoría de Respuesta al Ítem. *Memorias XXX Congreso Internacional de Investigación y Práctica Profesional en Psicología. Facultad de Psicología Universidad de Buenos Aires.*
- Weiss, D. J., y Sahin, A. (2024). *Computerized Adaptive Testing*. Guilford Publications.