

X Congreso Internacional de Investigación y Práctica Profesional en Psicología  
XXV Jornadas de Investigación XIV Encuentro de Investigadores en Psicología  
del MERCOSUR. Facultad de Psicología - Universidad de Buenos Aires, Buenos  
Aires, 2018.

## **Cómo hacer más breve un test breve: evaluación adaptativa del neuroticismo.**

Abal, Facundo Juan Pablo, Lozzia, Gabriela y  
Menéndez, Javier Agustín.

Cita:

Abal, Facundo Juan Pablo, Lozzia, Gabriela y Menéndez, Javier Agustín  
(2018). *Cómo hacer más breve un test breve: evaluación adaptativa del  
neuroticismo. X Congreso Internacional de Investigación y Práctica  
Profesional en Psicología XXV Jornadas de Investigación XIV Encuentro  
de Investigadores en Psicología del MERCOSUR. Facultad de Psicología -  
Universidad de Buenos Aires, Buenos Aires.*

Dirección estable: <https://www.aacademica.org/000-122/632>

ARK: <https://n2t.net/ark:/13683/ewym/e7D>

*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>.*

# CÓMO HACER MÁS BREVE UN TEST BREVE: EVALUACIÓN ADAPTATIVA DEL NEUROTICISMO

Abal, Facundo Juan Pablo; Lozzia, Gabriela; Menéndez, Javier Agustín

Universidad de Buenos Aires. Facultad de Psicología - Consejo Nacional de Investigaciones Científicas y Técnicas. Argentina

---

## RESUMEN

La versión breve del Cuestionario Revisado de Personalidad de Eysenck usa 12 ítems para medir la dimensión Neuroticismo. El objetivo de este trabajo es analizar si la administración adaptativa de los ítems permite una medición más eficiente del constructo. Participaron 594 personas de población general (52% mujeres). El 70% de los casos se usaron para calibrar los ítems con el Modelo Logístico de 2 parámetros de la Teoría de Respuesta al Ítem. El resto se empleó para simular la aplicación de un Test Adaptativo Informatizado (TAI) que interrumpe la evaluación al responder 6 ítems. Se encontró una correlación de .98 entre las estimaciones del rasgo obtenidas con el test completo y con el TAI. El error máximo registrado en una estimación (0.48) indica que ningún individuo fue evaluado con una confiabilidad clásica inferior a .77. El TAI tampoco afectó las evidencias de validez convergente y discriminante de la escala con las otras dimensiones del EPQ-RS y con las dimensiones sintomatológicas del SCL-90-R. Esto demuestra que la modalidad de evaluación adaptativa ensayada proporciona estimaciones de los niveles Neuroticismo de los evaluados con un grado de precisión aceptable aun cuando se administra 50% de los ítems de la escala del EPQ-RS.

## Palabras clave

Neuroticismo - TAI - Eysenck - EPQ-R

## ABSTRACT

### HOW TO MAKE MORE BRIEF A BRIEF TEST?: ADAPTIVE EVALUATION OF NEUROTICISM

The short version of the Eysenck Revised Personality Questionnaire (EPQ-RS) uses 12 items to measure the Neuroticism dimension. The objective of this work is to analyze if the adaptive administration of the items allows a more efficient measurement of the construct. 594 people from the general population participated (52% women). 70% of the cases were used to calibrate the items with the Logistic Model of 2 parameters of the Item Response Theory. The rest was used to simulate the application of a Computerized Adaptive Test (CAT) that interrupts the evaluation when answering 6 items. A correlation of .98 was found between the trait estimates obtained with the complete test and with the CAT. The maximum error recorded in an estimate (0.48) indicates that no individual was evaluated with a classic reliability inferior to .77. The CAT also did not affect the evidence of convergent and discriminant validity of the scale with the other dimensions of the EPQ-RS and with the symptomatic dimensions of the SCL-90-R. This shows that the modality of adaptive evaluation tested provides estimates of the Neuroticism levels of

those evaluated with an acceptable degree of precision even when 50% of the items on the EPQ-RS scale are administered.

## Keywords

Neuroticism - CAT - Eysenck - EPQ-R

El Neuroticismo es la tendencia de una persona a experimentar emociones negativas de manera intensa y/o por períodos inusualmente largos. Estas emociones suelen presentarse acompañadas por una marcada percepción de ineficacia para regularlas y para poder afrontar las situaciones cotidianas (Cassioello-Robbins, Wilner y Sauer-Zavala, 2017). Eysenck y Eysenck (1975) afirman que los sujetos con niveles altos de Neuroticismo son propensos a reaccionar de manera exagerada frente a cualquier tipo de estímulos y muestran dificultades para retornar al nivel emocional normal luego de estas experiencias. En cambio, las personas con niveles bajos de Neuroticismo se caracterizan por ser tranquilos, controlados y con un temperamento equilibrado en la mayoría de las situaciones. Eysenck (1947) fue el primero en identificar al Neuroticismo como una dimensión básica de la personalidad pasible de describir diferencias individuales y la incluyó en el nivel más alto de su modelo estructural junto con Extraversión y Psicoticismo. Entre las distintas dimensiones del modelo, el Neuroticismo es actualmente considerado prioritario en virtud del fuerte impacto que tiene en la salud pública (Lahey, 2009; Vittengl, 2017). Numerosos autores han comenzado a recomendar la detección de niveles altos de este dominio en población general durante la asistencia clínica de rutina (e.g. Allen y Kern, 2017; Hengartner, Kawohl, Haker, Rössler, y Ajdacic-Gross, 2016; Tackett y Lahey, 2017; Widiger, 2009) para implementar estrategias de intervención y prevención más tempranas y transdiagnósticas (Widiger y Oltmanns, 2017). La evaluación del Neuroticismo con fines de screening implica un nuevo desafío en la medición del constructo porque frente a estos objetivos son preferidos los instrumentos que, aun siendo muy breves, puedan ofrecer evidencias de calidad psicométrica.

La versión revisada del Cuestionario de Personalidad (EPQ-R) cuenta con una forma breve (EPQ-RS) en la que se utilizan 12 ítems para la medición del Neuroticismo (Eysenck, Eysenck, y Barrett, 1985). Esta parece ser una cantidad óptima de ítems porque consigue una adecuada representatividad del contenido y elevada confiabilidad. Algunos autores se propusieron acortar el EPQ-RS y ensayaron versiones abreviadas (EPQ-RA) de sólo 6 ítems (e.g. Francis, Brown, y Philipchalk, 1992; Sandín, Valiente, Chorot, Olmedo, y Santed, 2002). Pero la reducción en la cantidad de reactivos impactó negativamente en la confiabilidad de la escala, como es esperable

en el marco de la teoría clásica de test (Lewis, Francis, Shevlin, y Forrest, 2002).

El objetivo de este trabajo es estudiar si la implementación de un Test Adaptativo Informatizado (TAI) en el marco de la Teoría de Respuesta al Ítem - TRI - (Reise y Revicki, 2015) permite una reducción en la cantidad de ítems de Neuroticismo de la EPQ-RS mostrando ventajas sobre la administración tradicional.

## Método

### Participantes

Participaron 594 personas residentes en el área metropolitana de Buenos Aires, Argentina (52% género femenino). La edad osciló entre 18 y 82 años con una media de 32.9 años (DE = 12.1). El 79% consigné estar empleado y el 31% se encontraba cursando una carrera terciaria o universitaria al momento de ser evaluado. El 54.1% reportó tener un nivel de estudios secundarios completo y sólo un 6.2% no alcanzó este nivel.

### Instrumentos

*Escala Neuroticismo del Eysenck Personality Questionary revised short version, EPQ-RS* (Eysenck y Eysenck, 1985; adaptación de Squillace, Picón Janeiro y Schmidt, 2013). El cuestionario original consta de 48 ítems pero en la adaptación argentina quedó reducido a 42. Sólo 12 de estos ítems se corresponden a la dimensión Neuroticismo. La definición de esta dimensión establece que se conforma a partir de las intercorrelaciones de un conjunto de rasgos primarios: timidez, vergüenza, sentimiento de culpa, baja autoestima, mal humor, emotividad, depresión, miedo, ansiedad, preocupación, irracionalidad y tensión. Los adaptadores del instrumento replicaron la estructura factorial propuesta por Eysenck y registraron un índice de consistencia interna satisfactorio (KR-20 = .84) para la escala Neuroticismo.

*Inventario de síntomas SCL-90-R* (Derogatis, 1994; adaptación de Casullo y Perez, 2008). Consta de 90 ítems que se agrupan para posibilitar la medición de la intensidad de la sintomatología percibida durante la última semana en nueve dimensiones clínicas: Somatización, Obsesiones y Compulsiones, Sensitividad Interpersonal, Depresión, Ansiedad, Hostilidad, Ansiedad Fóbica, Ideación Paranoide y Psicoticismo. Además brinda tres índices globales: Índice de Severidad Global, Total de Síntomas Positivos e Índice de Malestar Positivo. La adaptación local muestra adecuadas evidencias de validez y estudios de confiabilidad.

### Procedimiento

Los datos se recolectaron a partir de evaluaciones individuales y sin tiempo límite. Los individuos debieron firmar un consentimiento en el que se explicitó el carácter voluntario de su participación y la posibilidad de abandonar la actividad si así lo deseaba. Se dieron garantías de anonimato y confidencialidad de sus respuestas.

### Análisis de datos

Se dividió a la muestra en dos grupos de manera aleatoria: a) un 70% (n=416) de los casos se utilizó para calibrar los ítems según el modelo logístico de dos parámetros (ML2p) y b) el 30% (n=178) fueron analizados para estudiar el comportamiento del TAI.

Previa aplicación de la TRI se corroboró el supuesto de unidimensionalidad requerido por el modelo con un Análisis Factorial Confirmatorio realizado con Mplus (Muthén y Muthén, 2010). La estimación de

los parámetros de los ítems se efectuó con el método de Máxima Verosimilitud Marginal operando el programa Multilog (Thissen, 2003). La bondad de ajuste del modelo se examinó ítem a ítem a partir de los métodos gráficos que brinda MODFIT (Stark, 2001).

Los ítems calibrados fueron usados como insumos para el estudio del TAI. Se llevó adelante una simulación sobre datos reales (simulación post hoc) operando el programa Firestar-D (Choi, 2009). Se usó la media del rasgo para la estimación inicial de  $\theta$ . Para las sucesivas estimaciones provisionales se utilizó el método bayesiano de estimación esperada a posteriori (EAP) considerando a la normal estándar como distribución a priori. La selección progresiva de los ítems se realizó según el método de máxima información de Fisher. Finalmente, se definió un criterio de parada de longitud fija por el cual la administración adaptativa finaliza al aplicar 6 ítems (correspondientes al 50% del total de la escala).

## Resultados

### Calibración de los ítems

Todos los índices de ajuste analizados permiten aceptar razonablemente la unidimensionalidad del Neuroticismo, TLI = .94, CFI = .95, RMSEA = .075, 90% IC [.063 - .087]. Las cargas factoriales fueron significativas ( $p < .05$ ) con valores que oscilaron entre .51 y .89.

La estimación de los parámetros del ML2p fue realizada en 36 ciclos hasta alcanzar un criterio de convergencia de 0.0001. El valor promedio de los parámetros de localización de los ítems ajustados fue de 0.24 (DE = 0.27, Mín = -0.04, Máx = 0.88).

Con relación a los parámetros de discriminación, el rango de los ítems ajustados varió entre 0.58 y 1.82. El promedio de  $a$  fue 1.15 (DE = 0.39), lo que supone que los ítems presentan una potencia discriminativa moderada en torno a sus respectivos valores de  $b$ . Mediante los gráficos de ajuste de MODFIT se constató que las curvas características de todos los ítems se mantuvieron dentro del intervalo de confianza asociado a la probabilidad observada para cada uno de los 25 niveles del rasgo contrastados. En conclusión, el ML2p resulta adecuado para modelizar los ítems de Neuroticismo.

### Estudio del TAI

Los niveles del rasgo estimados (simbolizados habitualmente con  $\theta$ ) con la versión adaptativa de 6 ítems correlacionaron de manera intensa y positiva con los niveles del rasgo estimados con el test completo,  $r = .97$ ,  $p < .001$ , y con el puntaje bruto clásico (suma de las puntuaciones de los ítems),  $r = .95$ ,  $p < .001$ .

Los errores en las estimaciones de  $\theta$  obtenidos al responder el TAI de longitud fija variaron entre 0.35 y 0.48 con una media de 0.39 (DE = 0.05). Este promedio supone que el conjunto de todos los sujetos fue medido con una confiabilidad clásica equivalente a .85. El error máximo registrado en la estimación (0.48) indica que ningún individuo fue evaluado con una confiabilidad inferior a 0.77.

Los porcentajes de uso de cada ítem en el TAI variaron ostensiblemente en función de las propiedades psicométricas de los elementos. En sintonía con el método de selección basado en la máxima información, los ítems más usados fueron aquellos con parámetros  $a$  más altos y parámetros  $b$  próximos a la media del rasgo. Por ejemplo, el ítem *¿Es usted inestable en sus estados de ánimo?* fue administrado en el 100% de los casos. En cambio, los ítems *¿Es usted una persona que se enoja con facilidad (irritable)?* y *¿Diría de*

*sí mismo que es una persona nerviosa?* no fueron seleccionados por el algoritmo para ninguno de los participantes debido a que presentaron los parámetros de discriminación más bajos.

Las evidencias de validez convergente y discriminante del Neuroticismo con el resto de las variables del EPQ-RS y las escalas del SCL-90-R no registraron diferencias importantes al comparar los coeficientes de correlación de las estimaciones surgidas del TAI o del test completo (tabla 1). Esto significa que la reducción en la cantidad de los ítems que presenta la versión adaptativa no impactó sustantivamente en las relaciones que el Neuroticismo guarda con estos constructos.

Tabla 1. Correlación de las estimaciones del TAI y el test completo con otros constructos.

	Neuroticismo (EPQ-RS)	
	TAI (6 ítems)	Test completo (12 ítems)
EPQ-RS		
Extraversión	.183**	.176**
Psicoticismo	.024	.032
Sinceridad	.127**	.118**
Escalas del SCL90-R		
Somatizaciones	.385**	.389**
Obsesiones y Compulsiones	.523**	.512**
Sensitividad interpersonal	.516**	.518**
Depresión	.656**	.673**
Ansiedad	.563**	.575**
Hostilidad	.404**	.395**
Ansiedad Fóbica	.396**	.388**
Ideación Paranoide	.321**	.348**
Psicoticismo	.467**	.443**
Índice de Severidad Global	.583**	.561**
Total de Síntomas Positivos	.456**	.473**
Índice de Malestar Positivo	.325**	.313**

Nota. \*\* $p < .001$

## Discusión

El interés de los investigadores por encontrar formas de medir rasgos de la personalidad con instrumentos breves resulta notorio cuando se observan las adaptaciones transculturales que ha tenido el EPQ-RS (e.g. Aluja, García, y García, 2003; Francis, Lewis, y Ziebertz, 2006; Lewis, et al., 2002; Squillace, et al., 2013; Tiwari, Singh, y Singh, 2009). Sin embargo, la teoría clásica presenta importantes limitaciones cuando se busca aumentar la eficiencia de las medidas ahorrando al máximo el tiempo de administración y conservando las propiedades psicométricas de las pruebas.

La TRI ha brindado las herramientas teórico-metodológicas necesarias para mejorar la medida del Neuroticismo de cara a su evaluación con fines de rastillaje. La modalidad de evaluación adaptativa ensayada en el presente trabajo ha demostrado que es posible proporcionar estimaciones de los niveles Neuroticismo de los evaluados con un grado de precisión entre aceptable y óptimo aun cuando se administran sólo el 50% de los ítems que componen la escala del EPQ-RS.

Resulta conveniente señalar que la reducción en la cantidad de ítems no ha perjudicado la representatividad del contenido muestreado (validez de contenido). Un TAI flexibiliza la evaluación de modo tal que cada sujeto ha respondido el subconjunto de seis ítems que más información aporta para su nivel de rasgo. La infra o supra representación de un aspecto específico constitutivo del Neuroticismo provocada por la medición adaptativa no tuvo consecuencias en la estimación de su rasgo. En efecto, la correlación elevada observada entre el  $\rho$  estimado con el TAI y  $\rho$  el obtenido con el test completo así lo demuestra. En la misma línea, tampoco se alteró de manera apreciable la asociación con otros constructos que forman parte de la red nomológica del Neuroticismo.

Los resultados de esta experiencia son alentadores porque confirman la posibilidad de evaluar el Neuroticismo de forma más rápida y precisa con los ítems del EPQ-RS. En un contexto de aplicación, el tiempo ahorrado sirve para evitar los efectos indeseados de la fatiga en la medida o, simplemente, podría ser aprovechado para la indagación de otras variables. En futuros estudios se buscará obtener evidencias de validez con respecto a un punto de corte que muestre relevancia clínica. Este punto abre la posibilidad de pensar un criterio de parada alternativo que apunte a administrar sólo los ítems relevantes para detectar los sujetos con niveles potencialmente clínicos en el rasgo.

## BIBLIOGRAFÍA

- Allen, K.A. y Kern, M.L. (2017). *School Belonging in Adolescents. Theory, Research and Practice*. Singapore: Springer.
- Aluja, A., García, O., y García, F.A. (2003). Psychometric analysis of the Revised Eysenck Personality Questionnaire Short Scale. *Personality and Individual Differences*, 35(2), 449-460. doi: 10.1016/S0191-8869(02)00206-4.
- Cassiello-Robbins, C., Wilner, J.G. y Sauer-Zavala, S. (2017). Neuroticism. En V. Zeigler-Hill y T.K. Shackelford (eds.). *Encyclopedia of Personality and Individual Differences*. Springer International Publishing. doi: 10.1007/978-3-319-28099-8\_1256-1
- Casullo, M.M. y Perez, M. (2008). *El Listado de Síntomas SCL-90-R de Derogatis*. Depto. de Publicaciones. Facultad de Psicología. UBA.
- Choi, S.W. (2009). Firestar: Computerized Adaptive Testing Simulation Program for Polytomous Item Response Theory Models. *Applied Psychological Measurement*, 33 (8), 644-645. doi: 10.1177/0146621608329892.
- Derogatis, L. (1994). *SCL-90-R. Symptom Checklist-90-R. Administration, Scoring and Procedures Manual*. Minneapolis: National Computer System.
- Dragow, F., Levine, M.V., Tsien, S., Williams B.A., y Mead, A.D. (1995). Fitting polytomous item response theory models to multiple-choice tests. *Applied Psychological Measurement*, 19(2), 143-165. doi: 10.1177/014662169501900203.
- Eysenck, H.J. (1947). *Dimensions of Personality*. Londres: Routledge y Kegan Paul.
- Eysenck, H.J., y Eysenck, S.B.G. (1985). *Manual of the Eysenck personality questionnaire*. London: Hodder and Stoughton.
- Eysenck, S.B.G., Eysenck, H.J., y Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences*, 6 (1), 21-29. doi: 10.1016/0191-8869(85)90026-1.
- Francis L.J., Lewis C.A, Ziebertz H.G. (2006). The Short-Form Revised Eysenck Personality Questionnaire (EPQR-S): A German edition. *Soc Behav Personal*, 34, 197-204. doi: 10.2224/sbp.2006.34.2.197.

- Francis, L.J., Brown, L.B., y Philipchalk, R. (1992). The development of an abbreviated form of the Revised Eysenck Personality Questionnaire (EPQR-A): its use among students in England, Canada, the USA and Australia. *Personality and Individual Differences*, 13 (4), 443-449. doi: 10.1016/0191-8869(92)90073-X.
- Hengartner, M.P., Kawohl, W., Haker, H., Rössler, W., y Ajdacic-Gross, V. (2016). Big Five personality traits may inform public health policy and preventive medicine: Evidence from a cross-sectional and a prospective longitudinal epidemiologic study in a Swiss community. *Journal of Psychosomatic Research*, 84, 44 - 51. doi: 10.1016/j.jpsychores.2016.03.012.
- Lahey, B.B. (2009). Public health significance of neuroticism. *American Psychologist*, 64, 241-256.
- Lewis, C.A., Francis, L.J., Shevlin, M., y Forrest, S. (2002). Confirmatory factor analysis of the French translation of the Abbreviated Form of the Revised Eysenck Personality Questionnaire (EPQR-A). *European Journal of Psychological Assessment*, 18, 179-185. doi: 10.1027//1015-5759.18.2.179.
- Muthén, L. y Muthén, B. (2010). *Mplus User's Guide, 6th Edn*. Los Angeles, CA: Muthén y Muthén.
- Reise, S.P. y Revicki, D.A. (2015). *Handbook of Item Response Theory Modeling Applications to Typical Performance Assessment*. Nueva York: Routledge.
- Sandín, B., Valiente, R.M., Chorot, P., Olmedo, M., y Santed, M.A. (2002). Versión española del cuestionario EPQR-Abreviado (EPQR-A) (I): Análisis exploratorio de la estructura factorial. *Revista de Psicopatología y Psicología Clínica*, 7(3), 195-205.
- Squillace, M., Picón Janeiro, J. y Schmidt, V. (2013). Adaptación local del Cuestionario Revisado de Personalidad de Eysenck (Versión abreviada). *Evaluar*, 13, 19 - 37.
- Stark, S. (2001). *MODFIT: A computer program for model-data fit*. Manuscrito no publicado. University of Illinois: Urbana-Champaign.
- Tackett, J.L. y Lahey, B.B. (2017). Neuroticism. En T.A. Widiger (Ed). *The Oxford handbook of the five factor model*. New York: Oxford University Press.
- Thissen, D. (2003). *MULTILOG*. Chicago: Scientific Software International.
- Tiwari, T., Singh, A.L., Singh, I.L. (2009). The Short-Form Revised Eysenck Personality Questionnaire: A Hindi edition (EPQRS-H). *Ind Psychiatry J*, 18 (1), 27-31. doi: 10.4103/0972-6748.57854.
- Vittengl, J.R. (2017). Who pays the price for high neuroticism? Moderators of longitudinal risks for depression and anxiety. *Psychological Medicine*, 1-12. doi:10.1017/S0033291717000253
- Widiger, T.A. y Oltmanns, J.R. (2017). Neuroticism is a fundamental domain of personality with enormous public health implications. *World Psychiatry*, 16 (2), 144-145. doi: 10.1002/wps.20411.
- Widiger, T.A. (2009). Neuroticism. En M.R. Leary y R.H. Hoyle (Eds), *Handbook of individual differences in social behavior* (pp. 129-146). New York; Guilford Press.