Avances en el estudio de la continuidad del neuroticismo y su variante desadaptativa: aplicación de la teoría de respuesta al ítem*

Advances in the Study of the Continuity of Neuroticism and its Maladaptive Variant: Application of the Item Response Theory

Juan Franco Sánchez González
ORCID: https://orcid.org/0000-0003-1993-4981
juansanchez31.31@gmail.com
Universidad de Buenos Aires, Argentina
Horacio Felix Attorresi
ORCID: https://orcid.org/0000-0003-2616-9202
horacioattorresi@gmail.com
Universidad de Buenos Aires, Argentina
Facundo Juan Pablo Abal
ORCID: https://orcid.org/0000-0001-7023-5380
Afjp79@gmail.co
Universidad de Buenos Aires, Argentina

Recibido: 20 septiembre 2020
Revisado: 07 marzo 2021
Aceptado: 02 septiembre 2021

Resumen

El Neuroticismo (N), a pesar de ser un rasgo de personalidad normal, se vincula con el padecimiento de diversas psicopatologias. Se estudió la continuidad entre el N y su variante desadaptativa, la Afectividad Negativa (AN), aplicando un modelo de la Teoría de Respuesta al Ítem (TRI) a dos escalas que miden estos constructos. Participaron 619 personas de población general (64.3 % mujeres), quienes respondieron un protocolo que incluyó 20 ítems de N del *International Personality Item Pool* (IPIP-NEO) y 23 ítems de AN del *Personality Inventory for DSM-5* (PID-5). Inicialmenete se realizó una depuración progresiva de los ítems de N y AN a partir de criterios psicométricos del Modelo de Respuesta Graduada. Luego se calibraron conjuntamente los 22 ítems retenidos aunque se eliminaron 6 ítems por mostrar dependencia local y problemas de ajuste. Se aportaron evidencias de la continuidad postulada entre N y AN. Sin embargo, se encontró un marcado solapamiento en la localización de los ítems procedentes de ambas escalas en el continuo del rasgo que conforman N y AN.

Palabras clave: neuroticismo, afectividad negativa, teoría de respuesta al ítem, personalidad, trastornos de personalidad.

Citar como: Sánchez González, J. F., Attorresi, H. F. y Abal, F. J. P. (2022). Avances en el estudio de la continuidad del neuroticismo y su variante desadaptativa: aplicación de la teoría de respuesta al ítem. *Diversitas: Perspectivas en Psicología, 18*(1). https://doi.org/10.15332/22563067.6125

Abstract

Neuroticism (N), despite being a normal personality trait, is linked to various psychopathologies. The continuity between N and its maladaptive variant, Negative Affectivity (NA), was studied by applying a model of the Item Response Theory (IRT) to two scales that measure these constructs. A total of 619 participants from the general population (64.3% women) answered a protocol that included 20 items of N from the International Personality Item Pool (IPIP-NEO) and 23 items of NA from the Personality Inventory for DSM-5 (PID-5). First, progressive refinement of the N and NA items was carried out based on psychometric criteria from the Graded Response Model. Then, the 22 items retained were calibrated together, although 6 items were eliminated for showing local dependence and adjustment problems. Evidence of the postulated continuity between N and NA was provided. However, a marked overlap was found in the location of the items from both scales in the trait continuum formed by N and AN.

Keywords: neuroticism, negative affectivity, item response theory, personality, personality disorders.

Introducción

El Neuroticismo (N) es un dominio que se encuentra ampliamente en la mayoría de modelos teóricos o empíricos de la estructura general de la personalidad normal. En el marco del Modelo de los Cinco Factores -*Five Factor Model*- (FFM, Costa y McCrae, 1995; Goldberg, 1993) se define como la tendencia del individuo a sentir afectos negativos de una manera intensa y prolongada. Así, personas con elevados niveles de N manifiestan preocupaciones frecuentemente, sentimientos de inseguridad, baja autoestima e insatisfacción consigo mismos y con el contexto; por lo que también tienden a establecer relaciones conflictivas. Resultan proclives a manifestar problemas en la autorregulación de emociones, lo que afecta de forma negativa a su capacidad de toma de decisiones y su manejo del estrés (McCrae y Costa, 2010). Algunos autores mencionan esta dimensión como Estabilidad Emocional, prefiriendo enfatizar el polo opuesto del dominio (Tackett y Lahey, 2017).

A pesar de que el N constituye un rasgo de personalidad normal, es abundante la evidencia empírica que muestra el impacto del N en el campo de la salud pública (Lahey, 2009; Vittengl, 2017; Widiger y Oltmanns, 2017). Su asociación con distintas psicopatologías ha llevado a concebirlo como un factor de vulnerabilidad común para el desarrollo y mantenimiento de diversos trastornos mentales (Barlow et ál., 2014; Gore y Widiger, 2018) y a diseñar tratamientos con objetivos centrados en la modificación de los niveles elevados de este rasgo (Drake et ál., 2017; Tackett y Lahey, 2017).

Las implicaciones del N en el sistema de salud pueden ser rastreadas por su vinculación con una gran variedad de enfermedades físicas como cardiopatías, diabetes, asma o enfermedades inflamatorias intestinales (Lahey, 2009). Niveles altos de N no solo son un factor de riesgo para la aparición de estas patologías, sino que también se ha observado un aumento en la tendencia a exagerar la importancia de los síntomas físicos y a obstaculizar la respuesta al tratamiento. Asimismo, el N se relaciona con el deterioro en la calidad de vida que repercute en las distintas áreas del individuo: problemas de rendimiento laboral o académico, dificultades interpersonales en los vínculos familiares, entre otros (Menéndez et ál., 2019; Vittengl, 2017).

Con la quinta versión del Manual de diagnóstico y estadística de los trastornos mentales (DSM-5; American Psychiatric Association [APA], 2013a) se ha generado un debate en el campo de la personalidad. Tomando como base el FFM se propuso un modelo dimensional

para los Trastornos de Personalidad - *Five Factor Model Personality Disorder* - (FFMPD, Clark, 2007, Widiger et ál., 2017; Widiger y Mullins-Sweatt, 2009). El FFMPD asume que para cada dominio de la personalidad normal es posible reconocer un polo desadaptativo o patológico. Esto incluye evidencia de una estructura jerárquica compartida entre rasgos normales y desadaptativos, al menos a nivel de los dominios (Al-Dajani et ál., 2015; Gore y Widiger, 2013; Widiger et ál., 2019).

Siguiendo el FFMPD, la APA (2013a) definió una variante desadaptativa del N llamada Afectividad Negativa (AN). Conceptualmente, este dominio guarda estrecha relación con el N, aunque con un mayor nivel de intensidad. La AN se caracteriza por la tendencia a vivenciar experiencias frecuentes e intensas de una amplia gama de emociones negativas; tales como la ansiedad, depresión, sentimiento de culpa o vergüenza, preocupación y enfado. Sus manifestaciones a nivel comportamental incluyen la autolesión y, en lo que respecta a lo interpersonal, se destaca la dependencia como forma de vincularse con otros (APA, 2013a).

La hipótesis que subyace a la nueva propuesta del DSM-5 supone la existencia de un continuo entre N y AN aunque establece que el límite entre ambos sería difuso y flexible (Sánchez y Ledesma, 2013). Los esfuerzos actuales se han centrado en aportar evidencias robustas sobre esta continuidad estudiando la relación de los diferentes instrumentos desarrollados para la medición del N y su variante desadaptativa. Al respecto, en la literatura se reconocen diversos abordajes metodológicos que apuntan a verificar este supuesto utilizando un enfoque correlacional (por ejemplo, Griffin y Samuel, 2014; Quilty et ál., 2013) o uno factorial (Wright y Simms, 2014). Sin embargo, estas estrategias metodológicas no permiten concluir que los rasgos de personalidad anormal son simplemente rasgos normales exacerbados (Walton et ál., 2008).

La aplicación de la Teoría de Respuesta al Ítem (TRI) como soporte metodológico permite alcanzar una mayor profundidad en la comprensión de esta continuidad (Balsis et ál., 2017). Los modelos de la TRI ofrecen la ventaja de estimar los parámetros que caracterizan a los ítems en una misma escala aun cuando pertenecen a distintos instrumentos (Reise y Waller, 2009; Walton et ál., 2008). Así, es posible comparar la información que proporcionan los distintos *tests* de personalidad normal y patológica (Samuel et ál., 2013; Samuel et ál., 2010; Stepp et ál., 2012) con el fin de analizar si los comportamientos más extremos de AN representan los niveles exacerbados del continuo que conforman N / AN. Más específicamente, con la TRI es factible determinar si, como se espera, los ítems que miden rasgos normales tienen capacidad discriminativa máxima en niveles medios de la variable mientras que los ítems que evalúan rasgos disfuncionales aportan información considerable en el nivel extremo superior del continuo (Suzuki et ál., 2015; Walton et ál., 2008).

Si bien son sólidas las evidencias que demuestran la existencia de una estructura dimensional común entre los dominios de los rasgos de personalidad normal y patológica (Clark, 2007; Widiger et ál., 2017; Widiger y Mullins-Sweatt, 2009) este tipo de investigaciones son poco frecuentes en el contexto latinoamericano y están limitadas a instrumentos en inglés. Mucho más escasos son aún los estudios que persiguen indagar esta continuidad haciendo uso de la TRI (Suzuki et ál., 2015; Walton et ál., 2008). A la luz de estas consideraciones teóricas y metodológicas, en este trabajo se propone estudiar el supuesto de continuidad entre la dimensión de N y su variante desadaptativa según el FFMPD aplicando la TRI en escalas adaptadas al contexto local que miden cada uno de estos constructos.

Método

Participantes

Se contó con la participación de 619 personas residentes en el área metropolitana de Buenos Aires, Argentina (64.3 % género femenino). Se usó un muestreo no probabilístico por accesibilidad. La edad de los participantes osciló entre 18 y 77 años con una media de 35.30 años (DE = 12.62). El 76.3 % reportó estar empleado y el 46.5 % se encontraba estudiando al momento de ser evaluado. El 38.7 % de los individuos alcanzó un nivel de estudios superior (Terciario o Universitario), el 56.2 % reportó tener un nivel de estudios de secundario completo y sólo un 5 % no alcanzó este nivel.

Instrumentos

Cuestionario de datos sociodemográficos. Encuesta breve diseñada ad-hoc que indaga datos aspectos tales como género, edad, cantidad de hijos, lugar de nacimiento y residencia, situación laboral y nivel educativo.

Ítems de la Escala de Afectividad Negativa del Personality Inventory for DSM-5 (PID-5, American Psychological Association, 2013b; Krueger et ál., 2012, adaptación de Fernández Liporace y Castro Solano, 2015). Si bien la escala original identifica nueve facetas anidadas jerárquicamente en AN (APA, 2013b) estudios posteriores conservaron solo tres debido a que las restantes mostraron asociaciones cruzadas con otros dominios del FFMPD (Morey et ál., 2013; Watters et ál., 2019). En consecuencia, se incluyeron los 23 ítems que componen las facetas Ansiedad (9 ítems, alfa de Cronbach = 0.82), Labilidad Emocional (7 ítems, alfa de Cronbach = 0.77) e Inseguridad Separación (7 ítems, alfa de Cronbach = 0.77). La consistencia interna de la escala total calculada a partir de los datos de la presente investigación resultó elevada (alfa de Cronbach = 0.87).

Ítems de Neuroticismo del International Personality Item Pool (IPIP-NEO, Goldberg et ál., 2006, adaptación de Cupani et ál., 2014). La escala completa está conformada por 48 elementos organizados en seis facetas, aunque para los fines de esta investigación se utilizó la versión del instrumento compuesta por una selección de 20 ítems (alfa de Cronbach = 0.87) que permiten una medición unidimensional del constructo.

Procedimiento

Se diseñó un protocolo autoadministrable en el que se alternaron los ítems de N del IPIP-NEO y AN del PID-5. Con el objetivo de lograr una mayor coherencia, todos los ítems fueron respondidos con una escala Likert de cuatro opciones: *En desacuerdo, Ligeramente en desacuerdo, Ligeramente de acuerdo y De acuerdo*. Esta elección se debe a estudios previos que mostraron que cuatro opciones pueden resultar una cantidad óptima para garantizar un equilibrio entre el grado de ajuste del modelo de la TRI y la precisión de la medida (e. g. Abal et ál., 2018).

La aplicación de los cuestionarios se realizó de forma individual y en un entorno adecuado, es decir, un *setting* de evaluación. Se informó a los participantes sobre los propósitos de la investigación, siendo su colaboración voluntaria. Se garantizó en todo momento el anonimato y la confidencialidad, así como el bienestar de los examinados, quienes fueron advertidos sobre la posibilidad de abandonar la evaluación si así lo deseaban. Se notificó previamente sobre la imposibilidad de recibir una devolución individual de los resultados del protocolo.

Tanto el protocolo como el diseño de la investigación fueron aprobados por el Comité

de Conductas Responsables de la Facultad de Psicología de la Universidad de Buenos Aires, sede en la que desarrolló este proyecto.

Análisis de Datos

En una primera fase se efectuó una depuración progresiva por separado de los ítems de N y de AN contemplando los criterios de calidad psicométrica de la TRI: independencia local, unidimensionalidad y ajuste al modelo. Se eliminó uno de los ítems de cada uno de los pares señalados como localmente dependientes por el estadístico x^2_{LD}. Siguiendo la recomendación de Reise y Rodríguez (2016), se consideró $x^2_{LD} > 10$ para determinar la dependencia local de los elementos. La unidimensionalidad se verificó con un Análisis Factorial Exploratorio (AFE) a partir de la matriz de correlaciones policóricas. Se aplicó el método de mínimos cuadrados no ponderados (ULS) para la extracción de factores y la implementación óptima de análisis paralelo de Horn como criterio de retención. También se analizó la proximidad a la unidimensionalidad con los índices ECV (*Explained Common Variance*) y MIREAL (*Mean of Item Residual Absolute Loadings*). Se contemplaron como aceptables valores de ECV > .85 y MIREAL < .30 (Ferrando y Lorenzo-Seva, 2018).

Finalmente, se calibraron los ítems con el Modelo de Respuesta Graduada (MRG) de Samejima (2016). La elección del MRG por sobre otros modelos politómicos de la TRI se asienta en su frecuente aplicación exitosa en *tests* de rendimiento típico (Reise y Revicki, 2015). Los parámetros se estimaron con el método de Máxima Verosimilitud Marginal con un criterio de convergencia de .0001 y el ajuste se examinó ítem por ítem con el índice S-x2 (Kang y Chen, 2011). Adicionalmente, se examinó el ajuste global del MRG con el M2, aunque dada la sensibilidad de este estadístico frente a pequeños desajustes de los datos al modelo también se complementó con el RMSEA (Maydeu-Olivares, 2015). De manera similar al criterio de interpretación definido para los modelos de ecuaciones estructurales (Maydeu-Olivares et ál., 2011) el RMSEA refleja un ajuste óptimo cuando presenta valores iguales o inferiores a 0.05.

En una segunda fase se procedió a calibrar conjuntamente los ítems retenidos para AN y N. Esto garantizó que la estimación de los parámetros de los ítems se encontraba en la misma métrica. En esta etapa se aplicaron los mismos criterios de depuración contemplados previamente. En ambas fases el conjunto de datos fue analizado haciendo uso del software IRTPRO 4.2 (Cai et ál., 2011) a excepción del AFE para el cual se utilizó el programa Factor 10 (Ferrando y Lorenzo-Seva, 2017).

Además, se utilizaron pruebas U de Mann Whitney con el objetivo de analizar si los parámetros de los ítems de los instrumentos difieren en función de su procedencia. Para hacer comparaciones que representen globalmente la *zona de actuación* de cada ítem se consideraron: a) los parámetros b_2 de cada ítem, b) el valor promedio de los parámetros de un mismo ítem (b_{prom}) y c) las amplitudes entre los valores de los parámetros b (*i.e.* b_3-b_1). También se cotejaron las Funciones de Información de los ítems que componen cada instrumento a fin de estudiar para qué niveles del rasgo resultan más precisas las escalas.

Resultados

Fase 1: Depuración de los instrumentos

Con respecto a la escala de AN, la depuración progresiva llevó a la eliminación de 11 de los 23 ítems originales. Siete de estos elementos fueron descartados debido a que mostraron dependencia local con otros reactivos (x^2_{LD}

> 10). Otros dos ítems mostraron valores de b poco razonables (por encima de ± 3) y los restantes no lograron ajustar al MRG según el índice S-x^2 (p < 0.001). Por último, el estudio factorial (*KMO* = 0.88; Prueba Bartlett, X^2 = 1775.3; *gl* = 66; p < 0.001) permitió verificar la unidimensionalidad del constructo con los ítems retenidos con base en los resultados obtenidos en la implementación óptima del análisis paralelo de Horn, el ECV = 0.86 y el MIREAL = 0.26.

En cuanto a la depuración de la escala de N, se descartaron 10 de los 20 ítems. Siete de estos reactivos fueron descartados por mostrar dependencia local con otros elementos (x^2_{LD} > 10) además de presentar parámetros de umbral con valores que excedían el rango esperable. Los últimos tres restantes se eliminaron por desajustar al MRG según los valores del índice $S\text{-}x^2$ (p < 0.001). En el AFE realizado (*KMO* = 0.88; Prueba Bartlett, X^2 = 1669.5; *gl* = 45; p < 0.001), la implementación óptima del análisis paralelo de Horn sugirió la extracción de un único factor, en consonancia con los índices de proximidad a la unidimensionalidad (ECV = 0.86, MIREAL = 0.20).

A modo de conclusión, en esta fase se eliminó el 48 % de los ítems de AN y el 50 % de los ítems de N. Se corroboró que los ítems conservados para cada escala formaron constructos unidimensionales y no presentaron dependencia local. Se obtuvieron evidencias de ajuste al MRG ítem por ítem con el índice $S\text{-}x^2$ así como también se corroboró el ajuste global tanto para AN (M_2 = 1183.9, *gl* = 582, p < 0.001; *RMSEA* = 0.04) como para N (M_2 = 734.2, *gl* = 395, p < 0.001; *RMSEA* = 0.04).

Fase 2: Calibración conjunta de los instrumentos

Los 22 ítems retenidos en la depuración de cada escala por separados fueron calibrados conjuntamente. Este análisis demandó una nueva depuración, en principio, por registrar problemas de dependencia local en pares de ítems provenientes de las escalas del AN y N. En este sentido, debieron descartarse dos ítems de AN y un elemento de N. Asimismo, tres ítems fueron suprimidos por evidenciar problemas de ajuste al MRG, dos pertenecientes a AN y dos provenientes de N.

Tabla 1.
Estimación y ajuste del Modelo de Respuesta Graduada

	a (s.e.)	b_1 *(s.e.)*	b_2 *(s.e.)*	b_3 *(s.e.)*	$S\text{-}x^2$	*gl*	*p*
IPIP01	1.60 (0.13)	-0.54 (0.09)	0.40 (0.07)	1.63 (0.12)	76.9	83	0.667
IPIP02	1.69 (0.15)	0.45 (0.07)	1.27 (0.10)	2.51 (0.19)	64.0	70	0.681
IPIP03	1.83 (0.15)	0.01 (0.07)	0.94 (0.08)	1.82 (0.13)	86.0	77	0.225
IPIP04	1.68 (0.15)	0.15 (0.07)	1.27 (0.10)	2.32 (0.17)	66.5	75	0.749
IPIP05	1.58 (0.13)	-0.21 (0.08)	0.81 (0.08)	1.89 (0.14)	77.5	83	0.651
IPIP06	2.00 (0.16)	-0.42 (0.07)	0.49 (0.06)	1.39 (0.10)	79.3	77	0.407
IPIP07	1.58 (0.13)	-1.05 (0.11)	0.02 (0.07)	1.02 (0.09)	85.4	80	0.319
PID01	1.11 (0.11)	-0.31 (0.10)	0.83 (0.10)	2.04 (0.19)	83.7	91	0.693
PID02	1.09 (0.11)	-0.05 (0.09)	1.13 (0.12)	2.63 (0.25)	103.3	92	0.198
PID03	1.38 (0.13)	0.19 (0.08)	1.26 (0.11)	2.33 (0.19)	80.4	81	0.499
PID04	1.49 (0.14)	0.13 (0.07)	0.98 (0.09)	2.03 (0.16)	85.2	84	0.443
PID05	1.76 (0.14)	-0.43 (0.08)	0.54 (0.07)	1.65 (0.12)	68.4	79	0.798
PID06	1.54 (0.13)	-1.07 (0.11)	-0.07 (0.07)	1.23 (0.10)	76.7	81	0.616
PID07	1.14 (0.11)	-1.78 (0.18)	-0.60 (0.10)	0.78 (0.10)	106.8	83	0.040
PID08	0.93 (0.12)	1.17 (0.15)	2.68 (0.31)	4.21 (0.52)	80.7	73	0.252
PID09	0.81 (0.09)	-1.71 (0.22)	-0.26 (0.12)	1.33 (0.17)	127.3	94	0.013

Fuente: elaboración propia.

En suma, para esta fase de calibración fueron conservados 16 ítems, 9 de estos de AN y los 7 restantes de N. El AFE aplicado (*KMO* = 0.92; Prueba Bartlett, X^2 = 2881.1; *gl* = 120; p < 0.001) corroboró la unidimensionalidad del constructo (ECV = 0.87, MIREAL = 0.21, implementación óptima del análisis paralelo de Horn) y el x^2_{LD} de todos los pares resultó inferior a 10. Con relación al ajuste del MRG, los *S-x2* obtenidos para los ítems de AN y N mostraron un ajuste satisfactorio (tabla 1). Así también, el análisis de ajuste global registró un índice de adecuado (M_2 = 1799.8, *gl* = 1064, p < 0.001; *RMSEA* = 0.03).

La tabla 1 reporta los resultados de la calibración conjunta de los 16 ítems retenidos de AN y N. El hecho de que los ítems provenientes de distintos instrumentos sean calibrados conjuntamente habilita la comparación de sus parámetros en virtud de que fueron estimados en la misma métrica. La estimación de los 64 parámetros del MRG requirió de 50 iteraciones para alcanzar el criterio de convergencia de 0.0001. Por cada uno de los reactivos se estimó un parámetro de pendiente (*a*) y tres parámetros de umbral b (b_1, b_2 y b_3). En tal sentido, el parámetro de b1 es entendido como el valor de umbral (medido en la escala del rasgo) que separa a la categoría En desacuerdo de las categorías superiores (*Ligeramente en desacuerdo, Ligeramente de acuerdo y De acuerdo*). Es decir, es el mínimo valor del rasgo necesario que requiere el individuo para tener una probabilidad igual o mayor a 0.50 de contestar la opción *Ligeramente en desacuerdo* (o una superior) en lugar de responder *En desacuerdo*. El resto de parámetros b de un mismo ítem se interpretan considerando una segmentación acumulativa de la respuesta politómica.

Una inspección de los contenidos de los ítems conservados luego de la depuración muestra que ni la escala N ni la AN se han visto perjudicada su cobertura a pesar de haberse reducido considerablemente. La escala N propone la medición de un dominio unidimensional y ha quedado representada la variedad de emociones negativas que describen a este constructo. En el caso de AN, al comparar el contenido de los 9 ítems retenidos se observa que las subdimensiones de este dominio aparecen representadas en una proporción similar a la identificada en el constructo original (4 ítems de Labilidad Emocional, 3 ítems de Ansiedad y 2 ítems de Inseguridad/Separación). En conclusión, las tres subdimensiones de la AN aparecen representadas en forma relativamente equitativa tras la reducción del instrumento.

Comparación de los resultados

No se registraron diferencias al comparar las zonas de actuación de los ítems provenientes de estos instrumentos. Ni las comparaciones de los parámetros b_2 (U = 29, p = 0.791), ni la de los b_{prom} promedio de los reactivos (U = 31, p = 0.958) fueron estadísticamente significativas en las pruebas U de Mann Whitney. Esto implica que, globalmente, los ítems de AN no se localizaron en niveles del rasgo más elevados que los ítems de N.

Sin embargo, se encontraron diferencias estadísticamente significativas en las amplitudes entre los parámetros *b* (i.e. b_3-b_1) de un mismo ítem (U = 10, p = 0.023). Estos resultados evidenciaron entonces que los reactivos pertenecientes a la escala de AN tienden a cubrir una mayor amplitud en el espectro del rasgo que los ítems provenientes de N. Este hallazgo resulta más comprensible al comparar Funciones de Información (FI) obtenidas para cada uno de los conjuntos de reactivos (figura 1). Aunque ambas FI se maximizan en torno a un θ de 0.8, la FI de los ítems de AN alcanza un menor grado de información (4.40) y se mantiene constante en un intervalo más amplio del rasgo latente.

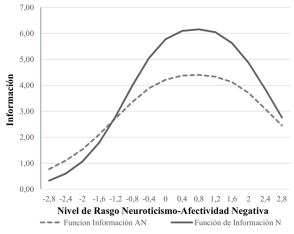

Figura 1.
Funciones de Información de los ítems de Neuroticismo y Afectividad Negativa
Fuente: elaboración propia.

Discusión

La continuidad entre rasgos normales y patológicos es un tema que en la literatura actual ha retomado su relevancia a partir del nuevo modelo dimensional propuesto por el DSM-5 para el estudio de los trastornos de la personalidad. En esta línea el presente trabajo permite aportar evidencia sobre esta hipótesis de continuidad, específicamente respecto al dominio de N y su variante desadaptativa. Al mismo tiempo, contribuye a generar conocimiento sobre esta temática en la población hispanohablante de Latinoamérica.

La depuración realizada en la primera fase permitió garantizar la calidad psicométrica con la que cada una de las escalas mide el constructo para el que fue construida. Se eliminó una gran cantidad de ítems como consecuencia de la aplicación del MRG. Esto se debe a que los supuestos que requiere la TRI tienen un mayor nivel de exigencia que los criterios que son usados por la Teoría Clásica de Test (Muñiz, 2018). Particularmente es común la violación del supuesto de dependencia local en escalas como el IPIP o el PID-5 que incluyen una gran cantidad de ítems redundantes para favorecer el aumento de la consistencia interna (Abal et ál., 2019; Abal et ál., 2010; Reise y Revicki, 2015; Reise y Rodríguez, 2016).

La calibración conjunta de los ítems de las dos escalas con TRI brindó resultados relevantes en dos sentidos. En primer lugar, permitió corroborar la continuidad hipotetizada entre N y AN, en tanto rasgos definidos desde una perspectiva normal y patológica. En segundo lugar, al tomar en consideración la ubicación de los ítems de ambos constructos en un mismo continuo, ofrece información sobre las zonas de actuación de los reactivos de cada instrumento (Krueger et ál., 2011; Reise y Waller, 2009; Walton et ál., 2008).

Si se toma en cuenta que el PID-5 fue construido para medir rasgos desadaptativos en población clínica, cabría esperar que sus reactivos midieran en distintos niveles del rasgo respecto de los que operacionalizan N con el IPIP, el cual se elaboró para población general. Es decir, los ítems del IPIP deberían ubicarse en los niveles medios del continuo que conforman N y AN. No obstante, los resultados de la comparación de los ítems en función de su instrumento de procedencia no reflejaron esta hipótesis. Por el contrario, es notorio el solapamiento encontrado en la ubicación de los ítems provenientes de cada técnica. Esto significa que los ítems de IPIP y PID-5 discriminan en similares niveles del rasgo latente. Consistentemente en diversos estudios se alcanzaron conclusiones idénticas a partir de la comparación de otros instrumentos que también miden rasgos normales y desadaptativos (Samuel et ál., 2010; Suzuki et ál., 2015; Walton et ál., 2008; Widiger y Costa, 2012).

Una posible explicación para este solapamiento se asienta en la selección realizada de las facetas de AN que son representadas en el PID-5. Es considerable la falta de consenso en la formulación teórica de las facetas que componen el dominio de AN (Watters y Bagby, 2018). Facetas como Tendencia depresiva, Desconfianza, Afectividad restringida y Hostilidad tienen una vinculación teórica relevante con la AN, pero fueron desestimadas por sus asociaciones con otros dominios desadaptativos del FFMPD (Watters et ál., 2019). Esta delimitación podría haber generado la pérdida de aquellos indicadores más extremos y, en consecuencia, que la versión de AN del PID-5 no esté incluyendo ítems relevantes de los niveles desadaptativos del rasgo para la población local. En efecto, los estudios que usaron instrumentos con contenidos más extremos, como la ideación suicida o la autolesión, sí reportaron diferencias en los niveles en que

resultan más precisas cada una de las escalas comparadas (Samuel et ál., 2010; Samuel et ál., 2013). En este sentido, sería conveniente revisar la operacionalización de las facetas que componen la dimensión de AN en la PID-5.

Otra posible interpretación del solapamiento registrado resulta de considerar que los ítems del IPIP estén proporcionando mayor información para la medición de niveles patológicos del rasgo, los cuales corresponderían al PID-5. Como concluyeron Ormel, Riese y Rosmalen (2012), los ítems con que habitualmente se mide N representan el mismo contenido que los instrumentos usados para evaluar sintomatología psicológica, por lo que la diferenciación entre los constructos parece borrosa. A esto se suma que la calibración conjunta realizada en este trabajo derivó en la eliminación de numerosos ítems que presentaron dependencia local, lo que refleja la fuerte semejanza entre los indicadores usados para la medición de N y AN.

Un punto importante para resaltar es la incidencia que podría tener las características inherentes del constructo que conforman N y AN. Diversas investigaciones reportaron dificultades a la hora de encontrar indicadores que logren representar el extremo inferior del dominio y sugieren que, en realidad, podría tratarse de un cuasi-rasgo y no de un constructo con dos polos como se plantea teóricamente (Abal et ál., 2019; Abal et ál., 2019; Gutiérrez et ál., 2017). Tanto las pruebas que miden N (Abal et ál., 2019) como AN (Gutiérrez et ál., 2017; McCrae et ál., 2005) fueron más precisas en la medición de niveles elevados del rasgo dado que registraron variaciones significativas para los puntajes altos, mientras que los puntajes bajos solo reflejan la ausencia de manifestaciones de este (Reise y Waller, 2009). El hecho de no alcanzar una correcta delimitación del polo inferior del constructo (Estabilidad Emocional) podría redundar en una dificultad para reconocer cuáles constituyen los indicadores comportamentales que caracterizarían los niveles medios de N y, consecuentemente, identificar conductas más extremas para definir AN.

El análisis realizado se ha centrado exclusivamente en el estudio de la continuidad entre N y AN aunque el modelo FFM y su variante patológica (FFMPD) incluyen otros cuatro dominios más. Los resultados aquí presentados respecto al solapamiento de los ítems en N y AN podrían no ocurrir en los otros rasgos que componen este modelo. En efecto, otros autores han encontrado diferencias en la localización de los ítems al comparar las versiones normales y desadaptativas de las dimensiones Agradabilidad y Responsabilidad (Samuel et ál., 2010). La decisión de priorizar N por sobre los demás dominios de la personalidad se basa en que ha sido señalado como un constructo clave para la descripción de la personalidad en el ámbito clínico y son justamente estos niveles elevados los que cobran relevancia por su vinculación con trastornos mentales y físicos (Brandes y Tackett, 2019; Jeronimus et ál., 2016; Widiger et ál., 2019). En este sentido, durante los últimos años se ha acumulado evidencia robusta de la asociación existente entre el N y un factor p o factor general de psicopatología propuesto en el modelo de Taxonomía Jerárquica de la psicopatología (*Hierarchical Taxonomy of Psychopathology*) (Brandes y Tackett, 2019; Kotov et ál., 2017).

En conclusión, los resultados de este estudio corroboran la continuidad entre N y AN, pero parece evidente que los límites entre ambos constructos resultan tan difusos que se presentan indistinguibles. Ahora bien, no es posible determinar si este fenómeno es inherente a las construcciones teóricas subyacentes o a sus respectivas operacionalizaciones. En sus formatos actuales, la escala AN del PID-5

y la escala N del IPIP podrían ser usadas indistintamente para la medición del gradiente entre lo normal y patológico del rasgo.

En cuanto a las limitaciones del presente estudio, en el plano metodológico cabe mencionar que los análisis se realizaron a partir de una muestra de población general. Futuros estudios se encaminarán en la aplicación de los ítems de ambos instrumentos también en población clínica con el fin de evaluar si existe un funcionamiento diferencial de los ítems al ser respondidos por ambos grupos. Asimismo, para descartar el potencial efecto de los instrumentos en las conclusiones sobre la continuidad de los constructos se considerará la inclusión de los ítems pertenecientes a las facetas no contempladas para la operacionalización de la AN con el PID-5. Esto podría ampliar la cantidad de reactivos de AN que se localizan en el extremo superior del continuo. Asimismo, convendría considerar la inclusión de ítems de otros instrumentos para la operacionalización de N, como el NEO-PI-R (McCrae y Costa, 2010).

Se espera que la consecución de esta línea de investigación no sólo contribuya a clarificar las relaciones existentes entre el N y su variante desadaptativa, sino además se desarrollen pruebas más sensibles para la detección de pacientes que presenten niveles de N / AN clínicamente relevantes.

Referencias

Abal, F. J. P., Auné, S. E. y Attorresi, H. F. (2018). Variación de la escala Likert en el test de utilidad de la matemática. *Interacciones, 4*(3), 177-189. https://doi.org/10.24016/2018.v4n3.134

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. Psicología, Cultura y Sociedad, 1*(1), 31-50. https://doi.org/10.18682/pd.v1i1.854

Abal, F. J. P., Lozzia, G. S., Auné, S. E. y Attorresi, H. F. (2019). Evaluación adaptativa del Neuroticismo mediante el Cuestionario Revisado de Personalidad de Eysenck. *Revista Evaluar, 19*(1). https://doi.org/10.35670/1667-4545.v19.n1.23877

Abal, F. J. P., Lozzia, G. S., Aguerri, M. E., Galibert, M. S., y Attorresi, H. F. (2010). La escasa aplicación de la teoría de respuesta al ítem en tests de ejecución típica. *Revista Colombiana de Psicología, 19*(1), 111-122.

Abal, F. J. P., Menéndez, J. A., y Attorresi, H. F. (2019). Aplicación de la Teoría de Respuesta al Ítem a la escala Neuroticismo del Big Five Inventory. *Ajayu Órgano de Difusión Científica del Departamento de Psicología UCBSP, 17*(2), 424-443.

Al-Dajani, N., Gralnick, T. M. y Bagby, R. M. (2015). A Psychometric Review of the Personality Inventory for DSM–5 (PID–5): Current Status and Future Directions. *Journal of Personality Assessment, 98*(1), 62–81. https://doi.org/10.1080/00223891.2015.1107572

American Psychiatric Association. (2013a). *Diagnostic and statistical manual of mental disorders* (5.ª ed.). American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2013b). *The Personality Inventory DSM-5 (PID-5) Self-Report Form (full version)*. http://www.psychiatry.org/practice/dsm/dsm5/online-assessmentmeasures#Personality

Balsis, S., Ruchensky, J. R. y Busch, A. J. (2017). Item response theory applications in personality disorder research. *Personality Disorders: Theory, Research, and Treatment, 8*(4), 298-308. https://doi.org/10.1037/per0000209

Barlow, D. H., Ellard, K. K., Sauer-Zavala, S., Bullis, J. R. y Carl, J. R. (2014). The Origins of Neuroticism. *Perspectives on Psychological Science, 9*(5) 481-496. https://doi.org/10.1177/1745691614544528

Brandes, C. M. y Tackett, J. L. (2019). Contextualizing neuroticism in the Hierarchical Taxonomy of *Journal of Research in Personality, 89*, 238-245. https://doi.org/10.1016/j.jrp.2019.06.007

Cai, L., Thissen, D. y Du Toit, S. (2011). *IRTPRO users guide*. Scientific Software International.

Clark, L. A. (2007). Assessment and diagnosis of personality disorder: Perennial issues and an emerging reconceptualization. *Annual Review of Psychology, 58*, 227-257. https://doi.org/10.1146/annurev.psych.57.102904.190200

Costa, Jr, P. T. y McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of personality assessment, 64*(1), 21-50. https://doi.org/10.1207/s15327752jpa6401_2

Cupani, M., Pilatti, A., Urrizaga, A., Chincolla, A. y Richaud, M. C. (2014). Inventario de personalidad IPIP-NEO: estudios preliminares de adaptación al español en estudiantes argentinos. *Revista Mexicana de Investigación en Psicología, 6*(1), 55-73. https://hdl.handle.net/11336/100851

Drake, M. M., Morris, D. M. y Davis, T. J. (2017). Neuroticism's susceptibility to distress: Moderated with mindfulness. *Personality and Individual Differences, 106*, 248–252. https://doi.org/10.1016/j.paid.2016.10.060

Fernández Liporace, M. L. y Castro Solano, A. (2015). Personality Inventory for DSM5. Adult Form. Argentinean Version. *Unpublished manuscript*.

Ferrando, P. J. y Lorenzo-Seva, U. (2017). Program FACTOR at 10: origins, development and future directions. *Psicothema, 29*(2), 236-241. https://doi.org/10.7334/psicothema2016.304

Ferrando, P. J. y Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement, 78*(5), 762-780. https://doi.org/10.1177/0013164417719308

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American psychologist, 48*(1), 26. https://doi.org/10.1037/0003-066X.48.1.26

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R. y Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in personality, 40*(1), 84-96. https://doi.org/10.1016/j.jrp.2005.08.007

Gore, W. L. y Widiger, T. A. (2013). The DSM-5 dimensional trait model and five-factor models of general personality. *Journal of Abnormal Psychology, 122*, 816-821. https://doi.org/10.1037/a0032822

Gore, W. L. y Widiger, T. A. (2018). Negative emotionality across diagnostic models: RDoC, DSM-5 Section III, and FFM. *Personality Disorders: Theory, Research, and Treatment, 9*(2), 155. https://doi.org/10.1037/per0000273

Griffin, S. A. y Samuel, D. B. (2014). A closer look at the lower-order structure of the Personality Inventory for DSM–5: Comparison with the five-factor model. *Personality Disorders: Theory, Research, and Treatment, 5*, 406. https://doi.org/10.1037/per0000074

Gutiérrez, F., Aluja, A., Peri, J. M., Calvo, N., Ferrer, M., Baillés, E., Gutiérrez-Zotes, J. A., Gárriz, M., Caseras, X., Markon, K. E. y Krueger, R. F. (2017). Psychometric properties of the Spanish

PID-5 in a Clinical and a Community Sample. *Assessment, 24*(3), 326-336. https://doi.org/10.1177/1073191115606518

Jeronimus, B. F., Kotov, R., Riese, H. y Ormel, J. (2016). Neuroticism's prospective association with mental disorders halves after adjustment for baseline symptoms and psychiatric history, but the adjusted association hardly decays with time: a meta-analysis on 59 longitudinal/prospective studies with 443 313 participants. *Psychological Medicine, 46*(14), 2883–2906. https://doi.org/10.1017/S0033291716001653

Kang, T. y Chen, T. T. (2011). Performance of the generalized S-X2 item fit index for the graded response model. *Asia Pacific Education Review, 12*(1), 89–96. https://doi.org/10.1007/s12564-010-9082-4

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R. Bagby, R. M., Brown, T. A. Carpenter, W. T., Caspi, A., Clark, L., Anna Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K. … Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454–477. https://doi.org/10.1037/abn0000258

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D. y Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*, 1879-1890. https://doi.org/10.1017/S0033291711002674

Krueger, R. F., Eaton, N. R., Clark, L. A., Watson, D., Markon, K. E., Derringer, J., Skodol, A. E. y Livesley, W. J. (2011). Deriving an empirical structure of personality pathology for DSM-5. *Journal of personality disorders, 25*(2), 170-191. https://doi.org/10.1521/pedi.2011.25.2.170

Lahey, B. B. (2009). Public health significance of neuroticism. *American Psychologist, 64*(4), 241. https://doi.org/10.1037/a0015309

Maydeu-Olivares, A. (2015). "Evaluating the Fit of IRT Models." In S Reise, D. Revicki (eds.), *Handbook of Item Response Theory Model: Applications to Typical Performance Assessment,* pp. 111–127. Routledge.

Maydeu-Olivares, A., Cai, L. y Hernández, A. (2011). Comparing the fit of item response theory and factor analysis models. *Structural Equation Modeling: A Multidisciplinary Journal, 18*(3), 333-356. https://doi.org/10.1080/10705511.2011.581993

McCrae, R. R., Löckenhoff, C. E. y Costa Jr., P. T. (2005). A step toward DSM-V: Cataloguing personality-related problems in living. *European Journal of Personality: Published for the European Association of Personality Psychology, 19*(4), 269-286. https://doi.org/10.1002/per.564

McCrae, R. y Costa P. (2010). *NEO Inventories professional manual*. Psychological Assessment Resources.

Menéndez, J. A., Attorresi, H. F, y Abal, F. J. P. (2019). Aportes de la medición del Neuroticismo en los ámbitos clínico y de la salud. *Anuario de Investigaciones, 26*.

Morey, L. C., Krueger, R. F. y Skodol, A. E. (2013). The hierarchical structure of clinician ratings of proposed DSM–5 pathological personality traits. *Journal of abnormal psychology, 122*(3), 836. https://doi.org/10.1037/a0034003

Muñiz, J. (2018). *Introducción a la psicometría*. Difusora Larousse-Ediciones Pirámide.

Ormel, J., Riese, H. y Rosmalen, J. G. (2012). Interpreting neuroticism scores across the adult life course: immutable or

experience-dependent set points of negative affect? *Clinical Psychology Review, 32*(1), 71-79. https://doi.org/10.1016/j.cpr.2011.10.004

Quilty, L. C., Ayearst, L., Chmielewski, M., Pollock, B. G. y Bagby, R. M. (2013). The psychometric properties of the Personality Inventory for DSM-5 in an APA DSM-5 Field Trial sample. *Assessment, 20*(3), 362-369. https://doi.org/10.1177/1073191113486183

Reise, S. P. y Revicki, D. A. (2015). *Handbook of Item Response Theory Modeling Applications to Typical Performance Assessment*. Routledge. https://doi.org/10.4324/9781315736013

Reise, S. P. y Rodriguez, A. (2016). Item response theory and the measurement of psychiatric constructs: some empirical and conceptual issues and challenges. *Psychological Medicine, 46*(10), 2025-2039. https://doi.org/10.1017/S0033291716000520

Reise, S. P. y Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology, 5*, 27-48. https://doi.org/10.1146/annurev.clinpsy.032408.153553

Samejima, F. (2016). Graded Response Model. En W. J. van der Linden (Ed.). *Handbook of Item Response Theory, Volume* 1: *Models* (pp. 95-108). Chapman y Hall/CRC.

Samuel, D. B., Carroll, K. M., Rounsaville, B. J. y Ball, S. A. (2013). Personality disorders as maladaptive, extreme variants of normal personality: Borderline personality disorder and neuroticism in a substance using sample. *Journal of Personality Disorders, 27*, 625– 635. https://doi.org/10.1521/pedi.2013.27.5.625

Samuel, D. B., Simms, L. J., Clark, L. A., Livesley, W. J. y Widiger, T. A. (2010). *An item response theory integration of normal and abnormal personality scales. Personality Disorders: Theory, Research, and Treatment, 1*(1), 5–21. https://doi.org/10.1037/a0018136

Sánchez, R. O. y Ledesma, R. D. (2013). Listado de Adjetivos para Evaluar Personalidad: Propiedades y normas para una población argentina. *Revista Argentina de Clínica Psicológica, XXII*, 147–160.

Stepp, S. D. Yu, L., Miller, J. D., Hallquist, M. N., Trull, T. J. y Pilkonis, P. A. (2012). Integrating competing dimensional models of personality: Linking the SNAP, TCI, and NEO using Item Response Theory. *Personality Disorders: Theory, Research, and Treatment, 3*, 107–126. https://doi.org/10.1037/a0025905

Suzuki, T., Samuel, D. B., Pahlen, S. y Krueger, R. F. (2015). DSM–5 alternative personality disorder model traits as maladaptive extreme variants of the five-factor model: An item-response theory analysis. *Journal of Abnormal Psychology, 124*, 343–354. https://doi.org/10.1037/abn0000035

Tackett, J. L. y Lahey, B. B. (2017). Neuroticism. In T. A. Widiger (Ed.), *Oxford library of psychology. The Oxford handbook of the Five Factor Model* (pp. 39-56). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199352487.013.14

Vittengl, J. R. (2017). Who pays the price for high neuroticism? Moderators of longitudinal risks for depression and anxiety. *Psychological Medicine*, 1-12. https://doi.org/10.1017/S0033291717000253

Walton, K. E., Roberts, B. W., Krueger, R. F., Blonigen, D. M. y Hicks, B. M. (2008). Capturing Abnormal Personality with Normal Personality Inventories: An Item Response Theory Approach. *Journal of Personality*. https://doi.org/10.1111/j.1467-6494.2008.00533.x

Watters, C. A. y Bagby, R. M. (2018). A meta-analysis of the five-factor internal structure of the Personality Inventory for DSM–5. *Psychological Assessment, 30*(9), 1255. https://doi.org/10.1037/pas0000605

Watters, C. A., Sellbom, M., Uliaszek, A. A. y Bagby, R. M. (2019). Clarifying the interstitial nature of facets from the Personality Inventory for DSM–5 using the five-factor model of personality. *Personality Disorders: Theory, Research, and Treatment*. https://doi.org/10.1037/per0000327

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. https://doi.org/10.1002/wps.20411

Widiger, T. A. y Costa, P. T. (2012). Integrating normal and abnormal personality structure: the five-factor model. *Journal of Personality, 80*(6), 1471-1506. https://doi.org/10.1111/j.1467-6494.2012.00776.x

Widiger, T. A. y Mullins-Sweatt, S. N. (2009). Five-factor model of personality disorder: A proposal for DSM-V. *Annual Review of Clinical Psychology, 5*, 197-220. https://doi.org/10.1146/annurev.clinpsy.032408.153542

Widiger, T. A., Gore, W. L., Crego, C., Rojas, S. L. y Oltmanns, J. R. (2017). Five Factor Model and Personality Disorder. *Oxford Handbooks Online*. https://doi.org/10.1093/oxfordhb/9780199352487.013.4

Widiger, T. A., Sellbom, M., Chmielewski, M., Clark, L. A., DeYoung, C. G., Kotov, R., Krueger, R. F., Lynam, D. R., Miller, J. D., Mullins-Sweatt, S., Samuel, D. B., South, S. C., Tackett, J. L., Thomas, K. M., Watson, D. y Wright, A. G. (2019). Personality in a hierarchical model of psychopathology. *Clinical Psychological Science, 7*(1), 77-92. https://doi.org/10.1177/2167702618797105

Wright, A. G. C. y Simms, L. J. (2014). On the structure of personality disorder traits: Conjoint analyses of the CAT-PD, PID-5, and NEO-PI-3 trait models. *Personality Disorders: Theory, Research, and Treatment, 5*(1), 43-54. https://doi.org/10.1037/per0000037

Notas

* Artículo de investigación. Estudio financiado por la Universidad de Buenos Aires, Proyecto UBACyT (UBACyT2018 20020170100200BA y UBACyT 2020 20020190200156BA).