

UNIVERSIDAD DE COSTA RICA
FACULTAD DE CIENCIAS SOCIALES
ESCUELA DE PSICOLOGÍA

**TRABAJO FINAL DE GRADUACION PARA OPTAR POR EL GRADO DE
LICENCIATURA EN PSICOLOGIA**

TITULO:

**Creación de una prueba psicométrica de auto reporte para la
evaluación dimensional de la psicopatía en población forense y general**

SUSTENTANTES:

Emmanuel Artavia Picado

Carné B00684,

Jimmy Chinchilla Alvarado

Carné B21930

Comité Asesor:

Director: M.Sc. Carlos Saborío Valverde

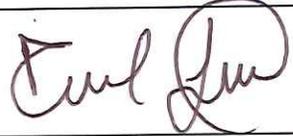
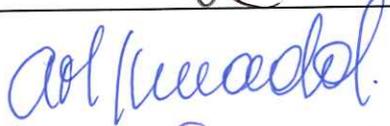
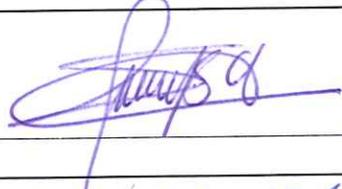
Lectora: Dra. Ana María Jurado Solórzano

Lectora: M.Sc. Shirley Barquero Gómez

2022

Nombre:

Firma:

Lic. José Manuel Mora Benamburg	
MPsc. Juan Carlos Miranda Orozco	
M.Sc. Carlos Saborío Valverde	
Dra. Ana María Jurado Solórzano	
M.Sc. Shirley Barquero Gómez	
Emmanuel Artavia Picado	
Jimmy Chinchilla Alvarado	

Resumen

Este proyecto final de graduación tuvo como objetivo la creación de un instrumento psicométrico para la evaluación de rasgos de personalidad psicopática de manera dimensional, y su validación en un contexto costarricense. Como marco teórico de referencia del constructo de psicopatía, y la creación del instrumento, se fundamentó en el modelo CAPP, Comprehensive Assessment of Psychopatic Personality (Cooke, et al., 2012).

Este modelo propone el abordaje de la psicopatía de manera dimensional, basado en síntomas característicos del trastorno de personalidad psicopática recopilados de diversos modelos explicativos sobre el constructo, estos fueron deconstruidos y definidos en términos de un lenguaje natural y descriptivo de rasgos. Los 33 síntomas que componen el CAPP se agruparon en seis dominios de personalidad y funcionamiento psicológico: Apego, Sí Mismo, Emocional, Conductual, Dominancia y Cognitivo.

Los ítems se construyeron con base en estos síntomas, bajo un modelo de escala Likert. Para su construcción, se tuvo en cuenta criterios de severidad, y la evaluación de conductas, actitudes, pensamientos y experiencias; estos se sometieron a una revisión exhaustiva mediante revisión con expertos y entrevistas cognitivas. Seguidamente, se realizó un pilotaje con grupo representativo de la muestra general y se llevó a cabo una revisión inicial de ítems mediante método Rasch para depurar el instrumento. Posteriormente, se aplicó la prueba a la muestra final; asimismo, se aplicaron pruebas adicionales para obtención de evidencias de validez de criterio: MMPI-2, IPPD, PAI, NEO PI-R y UPPS-P.

Para el análisis de datos, se hizo una aplicación de análisis estadísticos con el modelo de Rasch (Infit y Outfit); realización de análisis de funcionamiento diferencial de los ítems (DIF), Análisis de componentes principales de los residuos (ACPR), e igualmente, se llevaron a cabo análisis de alfa de Cronbach, análisis factorial exploratorio (AFE), análisis de T de Student, ANOVA, y coeficiente de Pearson.

De las 33 escalas, la mayoría logaron puntajes adecuados en los diferentes criterios estadísticos a los que se sometieron. No obstante, hubo algunas escalas que presentaron menor ajuste al modelo Rasch en criterios específicos de separación de personas. Esto está relacionado con la heterogeneidad de la muestra, lo cual se circunscribe en las limitaciones del proyecto y se

recomienda ampliar la muestra tanto general como forense para brindar datos más robustos de validación del instrumento.

Por otra parte, los resultados indican validez de criterio adecuados, dando evidencias de validez convergente y discriminante. Estos resultados, aportan evidencias positivas en tanto las escalas que evalúan el constructo a un nivel más amplio correlacionan con los diferentes dominios planteados a lo largo de la dimensión, lo que demuestra la capacidad de nuestro instrumento de captar diferentes síntomas a lo largo del continuo del constructo.

El instrumento creado demuestra una alta funcionalidad a lo largo de sus dominios y gran mayoría de sus escalas, y además ha logrado ajustarse en sus tres factores, alineándose las investigaciones más modernas del modelo CAPP. De esta manera, se concluye que se cumplieron a cabalidad los objetivos generales y específicos planteados para este trabajo final de graduación.

Palabras clave

Psicopatía, psicometría, Comprehensive Assesment of Psychopatic Personality, CAPP, PCL-R, dimensionalidad, Súper espectro de Externalización, personalidad, forense.

Referencia

Artavia Picado, E. y Chinchilla Alvarado, J. (2022). *Creación de una prueba psicométrica de auto reporte para la evaluación dimensional de la psicopatía en población forense y general*. [Tesis de licenciatura]. Universidad de Costa Rica, Costa Rica.

Agradecimientos y dedicatoria

A mi familia, especialmente a mi mamá por su apoyo incondicional.

A nuestro equipo asesor, Ana María y Shirley, por el apoyo y ayuda en este trabajo; a nuestro director Carlos, por alentarnos, apoyarnos y guiarnos a través de este ambicioso proyecto.

A mi amigo y compañero de tesis, Jimmy, por luchar conmigo hasta el final en esta aventura.

A José M. y Caro, especialmente agradecido, por su ayuda y colaboración con el proceso, sus aportes fueron invaluable para este trabajo.

A todos mis amigos y amigas que brindaron palabras de aliento y motivación, que me acompañaron y apoyaron incondicionalmente en este proceso.

A todas las personas que creyeron que lo lograríamos y nos apoyaron de una u otra forma.

A todas las personas que participaron como muestra de investigación y estuvieron dispuestas a colaborar, sin su ayuda este trabajo no hubiese sido posible.

A mí mismo, por mantener el rumbo a pesar de las vicisitudes de la vida, por ser fuerte y por no rendirme, por creer en mí.

- Emmanuel Artavia Picado

A mi papá, mi mamá y mis hermanas, por todo el amor y cariño demostrado en los momentos más difíciles.

Al equipo asesor por el apoyo brindado, al director de la tesis Carlos Saborío por proponernos este proyecto tan retador y enriquecedor. A mi compañero de tesis, y amigo, Emmanuel Artavia por atreverse a enfrentar este desafío conmigo.

Especial agradecimiento a José y Caro por sus importantes aportes que sin duda marcaron el resultado de este trabajo.

A todos mis amigos cercanos, quienes siempre estuvieron pendientes y apoyando de forma incondicional.

Infinitas gracias a cada uno de ustedes.

- Jimmy Chinchilla Alvarado

Tabla de contenidos

Agradecimiento y dedicatoria	I
Tabla de contenidos	II
Índice de tablas	IV
Índice de figuras	X
Introducción	1
Antecedentes de investigación	6
1. Antecedentes de investigación internacionales	7
2. Antecedentes nacionales	19
Marco teórico - conceptual	21
1. Trastornos de la personalidad: categóricos vs dimensionales	21
2. Críticas al Trastorno Antisocial de la Personalidad y DSM-5	24
3. Super Espectro de la Externalización: la psicopatía dentro de este modelo jerárquico	32
4. Un modelo dimensional de la psicopatía: CAPP	34
Problema de investigación	39
1. Objetivos generales	40
2. Objetivos específicos	40
3. Objetivo externo	41
Diseño metodológico	41
1. Tipo de estudio	41
2. Procedimiento general para la construcción de pruebas y su validación	42
3. Especificaciones del estudio	42
4. Descripción de la población y muestra	43
5. Confidencialidad y protección a los participantes	44
6. Retribución a la población	45
7. Sobre construcción, validación y pilotaje	45
8. Diseño y creación de ítems	46
9. Selección y revisión de ítems	47
10. Ensamble de la prueba	49
11. Desarrollar procedimientos y materiales para administrar y evaluar el test	49
12. Estudio piloto	51

13. Estudio de campo y aplicación de la prueba	52
14. Evidencias de validez de criterio	52
15. Análisis de datos	57
16. Resumen metodológico	63
Resultados	66
1. Análisis de Rasch	66
1.1. Dominio de Sí Mismo	67
1.2. Dominio Emocional	95
1.3. Dominio Dominancia	115
1.4. Dominio Apego	139
1.5. Dominio Cognitivo	156
1.6. Dominio Conductual	176
2. Evidencias de validez	199
Discusión de resultados	212
Conclusiones	226
Limitaciones	228
Recomendaciones	231
Referencias bibliográficas	233
Anexos	245

Índice de tablas

Tabla 1. Muestra respectiva para cada fase metodológica	46
Tabla 2. Resumen metodológico	63
Tabla 3: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Centrado en Sí Mismo	68
Tabla 4: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Centrado en Sí Mismo	68
Tabla 5: Estadísticos de los ítems para la escala Centrado en Sí Mismo	69
Tabla 6: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Autoagrandecimiento	72
Tabla 7: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Autoagrandecimiento	72
Tabla 8: Estadísticos de los ítems para la escala Autoagrandecimiento	73
Tabla 9: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Sentido de Ser Único	76
Tabla 10: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Sentido de Ser Único	76
Tabla 11: Estadísticos de los ítems para la escala Sentido de Ser Único	77
Tabla 12: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Sentido de Derecho	80
Tabla 13: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Sentido de Derecho	80
Tabla 14: Estadísticos de los ítems para la escala Sentido de Derecho	81
Tabla 15: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Sentido de Invulnerabilidad	84
Tabla 16: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Sentido de Invulnerabilidad	84
Tabla 17: Estadísticos de los ítems para la escala Sentido de invulnerabilidad	85
Tabla 18: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Autojustificación	88
Tabla 19: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Autojustificación	88

Tabla 20: Estadísticos de los ítems para la escala Autojustificación	89
Tabla 21: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Autoconcepto Inestable	92
Tabla 22: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Autoconcepto Inestable	93
Tabla 23: Estadísticos de los ítems para la escala Autoconcepto Inestable	93
Tabla 24: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Falta de Ansiedad	96
Tabla 25: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Falta de Ansiedad	97
Tabla 26: Estadísticos de los ítems para la escala Falta de Ansiedad	97
Tabla 27: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala No Experimenta Placer	100
Tabla 28: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala No Experimenta Placer	101
Tabla 29: Estadísticos de los ítems para la escala No Experimenta Placer	101
Tabla 30: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Falta de Profundidad Emocional	104
Tabla 31: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Falta de Profundidad Emocional	105
Tabla 32: Estadísticos de los ítems para la escala Falta de Profundidad Emocional ...	105
Tabla 33: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Falta de Estabilidad Emocional	108
Tabla 34: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Falta de Estabilidad Emocional	109
Tabla 35: Estadísticos de los ítems para la escala Falta de Estabilidad Emocional	109
Tabla 36: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Falta de Remordimiento	112
Tabla 37: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Falta de Remordimiento	113

Tabla 38: Estadísticos de los ítems para la escala Falta de Remordimiento	113
Tabla 39: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Antagónico	116
Tabla 40: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Antagónico	116
Tabla 41: Estadísticos de los ítems para la escala Antagónico	117
Tabla 42: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Dominante	120
Tabla 43: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Dominante	120
Tabla 44: Estadísticos de los ítems para la escala Dominante	121
Tabla 45: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Engañoso	124
Tabla 46: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Engañoso	124
Tabla 47: Estadísticos de los ítems para la escala Engañoso	125
Tabla 48: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Manipulador	128
Tabla 49: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Manipulador	128
Tabla 50: Estadísticos de los ítems para la escala Manipulador	129
Tabla 51: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Insincero	132
Tabla 52: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Insincero	132
Tabla 53: Estadísticos de los ítems para la escala Insincero	133
Tabla 54: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Charlatán	136
Tabla 55: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Charlatán	136
Tabla 56: Estadísticos de los ítems para la escala Charlatán	137
Tabla 57: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Desapegado	140

Tabla 58: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Rasgos Impulsivos	140
Tabla 59: Estadísticos de los ítems para la escala Rasgos Impulsivos	141
Tabla 60: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de No Adquiere Compromisos	144
Tabla 61: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de No Adquiere Compromisos	145
Tabla 62: Estadísticos de los ítems para la escala de No Adquiere Compromisos	146
Tabla 63: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Sin Empatía	149
Tabla 64: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Sin Empatía	150
Tabla 65: Estadísticos de los ítems para la escala Sin Empatía	150
Tabla 66: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Indiferente	153
Tabla 67: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Indiferente	153
Tabla 68: Estadísticos de los ítems para la escala de Indiferente	154
Tabla 69: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Suspicaaz	157
Tabla 70: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Suspicaaz	158
Tabla 71: Estadísticos de los ítems para la escala de Suspicaaz	158
Tabla 72: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Falta de Concentración	161
Tabla 73: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Falta de Concentración	162
Tabla 74: Estadísticos de los ítems para la escala de Falta de Concentración	162
Tabla 75: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Intolerante	165
Tabla 76: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Intolerante	166

Tabla 77: Estadísticos de los ítems para la escala de Intolerante	166
Tabla 78: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Inflexible	169
Tabla 79: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Inflexible	170
Tabla 80: Estadísticos de los ítems para la escala de Inflexible	170
Tabla 81: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Falta de Planeamiento	173
Tabla 82: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Falta de Planeamiento	174
Tabla 83: Estadísticos de los ítems para la escala de Falta de Planeamiento	174
Tabla 84: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Inconstante	177
Tabla 85: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Inconstante	178
Tabla 86: Estadísticos de los ítems para la escala de Inconstante	178
Tabla 87: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Poco Confiable	181
Tabla 88: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Poco Confiable	182
Tabla 89: Estadísticos de los ítems para la escala de Poco Confiable	182
Tabla 90: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Imprudente	185
Tabla 91: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Imprudente	186
Tabla 92: Estadísticos de los ítems para la escala de Imprudente	186
Tabla 93: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Intranquilo	189
Tabla 94: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Intranquilo	189
Tabla 95: Estadísticos de los ítems para la escala de Intranquilo	190
Tabla 96: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Disruptivo	193

Tabla 97: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Disruptivo	194
Tabla 98: Estadísticos de los ítems para la escala de Disruptivo	194
Tabla 99: Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Agresivo	197
Tabla 100: Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Agresivo	198
Tabla 101: Estadísticos de los ítems para la escala de Agresivo	198
Tabla 102: KMO y Test de esfericidad de Bartlett	200
Tabla 103: Varianza explicada por factores	200
Tabla 104: Matriz de estructura para las dimensiones	202
Tabla 105: Correlaciones de dominios con escalas del MMPI-2	203
Tabla 106: Correlaciones de dominios con escalas del PAI	204
Tabla 107: Correlaciones de dominios con escalas del UPPS-P	205
Tabla 108: Correlaciones de dominios con escalas del IDPP	206
Tabla 109: Correlaciones de dominios con escalas del NEO PI-R	207
Tabla 110: Tabla comparativa de dominios: Medias, desviaciones estándar, prueba t y significancia bilateral para la muestra según género	208
Tabla 111: Resultados ANOVA según cada dominio	209
Tabla 112: Comparaciones Post Hoc – Grupo de edad del dominio Conductual	210
Tabla 113: Comparaciones Post Hoc - Grupo de edad del dominio de Apego	210
Tabla 114: Comparaciones Post Hoc - Grupo de edad del dominio de Dominancia ..	211
Tabla 115: Comparaciones Post Hoc - Grupo de edad del dominio de Sí Mismo	212

Índice de figuras

Figura 1. Comprehensive Assessment of Psychopathic Personality (CAPP)	37
Figura 2. Gráfico Escala Centrado en Sí Mismo – Distribución casos e ítems	67
Figura 3: Gráfico Escala Autoagrandecimiento – Distribución casos e ítems	71
Figura 4: Gráfico Escala Sentido de Ser Único – Distribución casos e ítems	75
Figura 5: Gráfico Escala Sentido de Derecho – Distribución casos e ítems	79
Figura 6: Gráfico Escala Sentido de Invulnerabilidad – Distribución casos e ítems	83
Figura 7: Gráfico Escala Autojustificación – Distribución casos e ítems	87
Figura 8: Gráfico Escala Autoconcepto Inestable – Distribución casos e ítems	91
Figura 9: Gráfico Escala Falta de Ansiedad – Distribución casos e ítems	95
Figura 10: Gráfico Escala No Experimenta Placer – Distribución casos e ítems	99
Figura 11: Gráfico Escala Falta de Profundidad Emocional – Distribución casos e ítems	103
Figura 12: Gráfico Escala Falta de Estabilidad Emocional – Distribución casos e ítems	107
Figura 13: Gráfico Escala Falta de Remordimiento – Distribución casos e ítems	111
Figura 14: Gráfico Escala Antagónico – Distribución casos e ítems	115
Figura 15: Gráfico Escala Dominante – Distribución casos e ítems	119
Figura 16: Gráfico Escala Engañoso – Distribución casos e ítems	123
Figura 17: Gráfico Escala Manipulador – Distribución casos e ítems	127
Figura 18: Gráfico Escala Insincero – Distribución casos e ítems	131
Figura 19: Gráfico Escala Charlatán – Distribución casos e ítems	135
Figura 20: Gráfico de Escala Desapegado – Distribución casos e ítems	139
Figura 21: Gráfico Escala No Adquiere Compromisos – Distribución casos e ítems...	143
Figura 22: Gráfico Escala Sin Empatía – Distribución casos e ítems	148
Figura 23: Gráfico Escala de Indiferente – Distribución casos e ítems	152
Figura 24: Gráfico Escala Suspica – Distribución casos e ítems	156
Figura 25: Gráfico Escala de Falta de Concentración – Distribución casos e ítems	160
Figura 26: Gráfico Escala de Intolerante – Distribución casos e ítems	164
Figura 27: Gráfico Escala de Inflexible – Distribución casos e ítems	168
Figura 28: Gráfico Escala Falta de Planeamiento – Distribución casos e ítems	172
Figura 29: Gráfico Escala de Inconstante – Distribución casos e ítems	176

Figura 30: Gráfico Escala de Poco Confiable – Distribución casos e ítems	180
Figura 31: Gráfico Escala de Imprudente – Distribución casos e ítems	184
Figura 32: Gráfico Escala de Intranquilo – Distribución casos e ítems	188
Figura 33: Gráfico Escala Disruptivo – Distribución casos e ítems	192
Figura 34: Gráfico Escala de Agresivo – Distribución casos e ítems	196

Introducción

En la actualidad existe un debate respecto de la clasificación de la psicopatología, en tanto se considera tradicionalmente desde un enfoque categórico, para el cual cumplir ciertos criterios preestablecidos determina si una persona se diferencia cualitativamente de la normatividad (Patrick, 2006; DeLisi, 2016); sin embargo, en los últimos años se ha aportado evidencia que respalda un enfoque dimensional del abordaje de la psicopatología (Kotov et al., 2018; Krueger et al., 2021; Krueger et al., 2020).

La psicometría permite la evaluación de la psicopatología y su dimensionalidad de una manera científica; adicionalmente, propicia la creación de evidencia que respalda esta concepción dimensional de los constructos, esto mediante la validación del test, lo cual implica un juicio evaluativo global que garantiza apoyo a sus interpretaciones (Popham, 2000).

Asociado a lo anterior, la validez de constructo refiere al “grado en que las relaciones de las puntuaciones de los tests con otras conductas reflejan las relaciones esperadas en la teoría del constructo” (Martínez-Arias et al., 2014, p. 235), lo cual es de suma importancia, ya que permite diseñar pruebas apropiadas, que cuenten con evidencia de su validez y confiabilidad, para el estudio de constructos específicos en la psicopatología.

En las últimas décadas, el abordaje del constructo de la psicopatía ha contado con vasta investigación empírica, mucha de la cual ha girado en torno a los planteamientos de Hare y otros autores basados en el modelo del PCL-R (Patrick, 2006; DeLisi, 2016). Miller et al. (2016) definen la psicopatía como un trastorno de la personalidad que comprende características desadaptativas y que está asociado con problemáticas de comportamiento antisocial. Además, existe evidencia de que es un constructo dimensional (Guay et al., 2007) que se puede conceptualizar como una configuración de variaciones extremas o desadaptativas de características generales de la personalidad.

Hare en 1996 (como se citó en Guay et al., 2007) definió la psicopatía como un trastorno de personalidad que comprende un conjunto de características emocionales, interpersonales y conductuales, que se caracteriza por la indiferencia hacia las reglas sociales y los derechos de otros.

Estos postulados se fundamentan en la propuesta bifactorial de Hare (2003), en la que el constructo de la psicopatía se compone de un primer factor afectivo e interpersonal y un segundo factor de estilo de vida antisocial. El abordaje de la psicopatía que hace Hare tiene una gran influencia de lo planteado por Cleckley (1988), quien identificó 16 características o rasgos del trastorno psicopático, las cuales aún influyen la investigación actual sobre el tema (ver anexo 1).

No obstante, solamente cuatro de las 16 características concuerdan con las 20 características de psicopatía planteadas por Hare. De esta forma, existe una marcada influencia, pero la conceptualización de Hare caracteriza y delimita un poco más específicamente el constructo.

Ante el planteamiento de constructos que comprenden la psicopatía, Hare (2003) utiliza un abordaje de la medición para el estudio de esta, siendo este el siguiente paso científico en el estudio del trastorno. En general, identifica 20 características de la psicopatía, que a su vez se subdividen dentro de dos factores, 1. Interpersonal y afectivo, y 2. Estilo de vida antisocial, y 4 facetas: Interpersonal, Afectiva, Estilo de vida y Antisocial (ver anexo 2).

La propuesta de psicopatía planteada por Hare (2003), en el PCL-R, es el estandarte actual de la conceptualización de la psicopatía; de esta manera se toma como base la división en factores del constructo, donde el Factor 1 abarca mayormente las características de personalidad y el Factor 2 la parte comportamental antisocial del constructo psicopático.

Adicionalmente, es importante mencionar que el estudio de este constructo se realiza mayormente en población forense, privados de libertad e implicados en asuntos judiciales; el abordaje de esta condición en población comunitaria sin antecedentes penales/delictivos es escaso. Sin embargo, también se han desarrollado abordajes como el de Lilienfeld (1990; 2005) mediante la construcción de un instrumento de auto reporte, el *Psychopathic Personality Inventory* (PPI), desarrollado y validado con población no forense, lo que permite una evaluación más acertada en población no forense o clínica.

Ahora bien, la tendencia al comportamiento antisocial y sus características afectivas propias implica posibles riesgos en diferentes ámbitos sociales, por lo que una óptima evaluación de estas características es de mucha utilidad para establecer factores de riesgo o adaptación a situaciones específicas. Además, el estudio de los trastornos y condiciones

mentales, como es la personalidad psicopática, permite un mejor entendimiento de estos constructos y esto puede facilitar un mejor abordaje a futuro de las personas, en este caso, que tienen rasgos psicopáticos; llámese en tratamiento, con recursos de apoyo, en rehabilitación, entre otros (Salvador et al, 2017, citado en Saborío, 2012). Adicionalmente, estos mismos autores mencionan que los principales instrumentos y familias de referencia de la evaluación de la psicopatía son el *Psychopathy Checklist* (PCL), las *Levenson's Primary and Secondary Psychopathy Scales* (LPSP), el *Self-Report Psychopathy Scale* (SRP), y el *Psychopathic Personality Inventory* (PPI). Estos instrumentos son altamente reconocidos y utilizados a nivel internacional, especialmente el PCL, que se ha tomado como un máximo referente, por su modelo teórico de psicopatía.

En cuanto a Latinoamérica, existen adaptaciones del PCL-R tanto en Chile (Abalos et al., 2004), como en México (Ostrosky-Solís et al., 2008); sin embargo, hay una carencia de métodos de evaluación psicométricos para el constructo que sean originarios de la zona. Esta es una razón que resalta la importancia de crear un instrumento con el objetivo de evaluar características de personalidad psicopática directamente en una población hispanoparlante y de un contexto latinoamericano.

De manera específica, en Costa Rica, el estudio de la psicopatía es sumamente limitado, así como la creación o validación de herramientas psicométricas para el estudio de este trastorno. Existen algunas investigaciones, principalmente trabajos finales de graduación, que exploran la relación entre personas privadas de libertad y la psicopatía mediante la medición del PCL-R u otros instrumentos (Arce y Gutiérrez, 2003; Gamboa y Pérez, 2004; Masís, 2008; Mora y Purcallas, 2005; Picado, 2007; Pisoni, 2006; Reyes, 2006). Las pocas investigaciones científicas relacionadas con en el tema dejan entrever la necesidad de ahondar sobre el mismo en el contexto nacional, mediante el uso de la psicometría y las teorías de dimensionalidad y externalización. Ante esto, queda evidenciada la relevancia, pertinencia y necesidad de la construcción de un instrumento para la medición de características psicopáticas, a través de un estudio de validación, que se encuentre ajustado a las características de la población costarricense, y que además aporte datos normativos para la misma.

De esta forma, en el presente trabajo final de graduación se construyó un instrumento psicométrico para la evaluación de características de personalidad psicopática. La creación

de un *test* psicométrico cuya finalidad es evaluar rasgos de la psicopatía en Costa Rica, siendo pioneros en Centroamérica y Latinoamérica, puede ser un paso inicial para impulsar el estudio de este tipo de fenómenos en la región y producir herramientas más adecuadas (óptimas, confiables y válidas) ajustadas al contexto nacional e investigaciones sobre este trastorno en poblaciones específicas que no han sido abordadas.

En esta misma línea, construir un instrumento ajustado a la población costarricense, validado y con lenguaje ajustado a la población nacional permitirá una mejor evaluación del constructo en la misma población, así como tener una valoración más precisa y mejores indicadores en poblaciones con características similares. Siendo todo esto un aspecto muy importante para el fomento de una práctica profesional basada en evidencia en el país que permita una intervención a futuro más adecuada en este tipo de condiciones psicológicas. Para ello será fundamental apoyarse en la validez de constructo, la cual, según Leyva (2011), permite demostrar con evidencia, el grado en que los puntajes obtenidos en la prueba realmente representan los diferentes rasgos de psicopatía, los cuales se van a proponer más adelante.

Por otro lado, una de las limitaciones metodológicas que se ha resaltado en la literatura es que la gran mayoría del estudio de la psicopatología, y en particular de la psicopatía, se realiza desde un enfoque taxonómico o categórico (American Psychiatric Association, 2013; Wright, 2009; DeLisi, 2016), además, se ha evidenciado su poca utilidad y validez de constructo; esto ha hecho que se requiera un abordaje del trabajo a partir de un enfoque dimensional. En este sentido, instrumentos que brindan un abordaje dimensional del constructo, como el PPI (Lilienfeld, 1990), han representado un gran aporte al estudio del constructo.

Ligado a lo anterior, la población a la cual va dirigido el instrumento afecta en gran medida la capacidad de este para evaluar el constructo de manera dimensional. Una ventaja que brindaría la construcción de un instrumento dimensional del trastorno psicopático, es la capacidad para evaluar a diferentes poblaciones que se extiendan desde una persona socialmente funcional y sin antecedentes delictivos hasta una persona con severo y amplio historial delictivo. Esto debido a que actualmente no existen herramientas desarrolladas - empírica y conceptualmente- con el fin de evaluar la psicopatía en población general. A su

vez, la restricción de la evaluación en población forense puede limitar la dimensionalidad del constructo en tanto no abarca esta parte de la dimensión, a la población general o comunitaria.

Asimismo, un instrumento que sea aplicable a diferentes poblaciones permitirá mayor utilidad a nivel práctico y no estaría limitado al ámbito forense, como comúnmente sucede con medidas de evaluación de la personalidad psicopática. Entre estas aplicaciones se puede mencionar la selección de personal, en donde los psicólogos laborales y organizacionales y el personal de recursos humanos y reclutamiento podrían servirse de un instrumento de este tipo para identificar características deseables o no deseables de acuerdo con los puestos laborales y así determinar la idoneidad de la persona para el mismo. En este mismo punto, organizaciones gubernamentales de seguridad pública, fuerzas policiales y jurídicas, así como compañías de seguridad privada, podrían servirse de este instrumento para realizar una evaluación específica de sus empleados y oficiales, para identificar posibles factores de riesgo y así determinar la idoneidad para los diferentes puestos de acuerdo con las características del individuo y las del propio puesto de trabajo.

De forma similar, este instrumento podría ayudar a los profesionales en la evaluación de personas en los procesos propios de otorgación de permisos para la portación de armas o licencias de diferentes índoles. Dado que estos puestos de trabajo y permisos pueden implicar un riesgo tanto para la persona beneficiaria como para terceros, es de suma importancia realizar una evaluación adecuada para determinar si existen factores de riesgo en la persona, factores que podrían incluir rasgos de una personalidad psicopática; en cuyo caso la persona podría no ser idónea para desempeñar esa labor u obtener ese permiso o licencia.

Finalmente, el instrumento construido es de auto-reporte, lo que presenta muchos beneficios por encima de la medición tradicional basada en la entrevista. Lilienfeld y Fowler (2006) indican varios puntos: 1) al ser la misma persona la observadora permite la evaluación de estados y características emocionales subjetivas, o en el caso de la psicopatía, la ausencia de estas. 2) Economía, ya que es más corto y de fácil administración; evaluación de estilos de respuesta como manejo de la impresión y simulación que pueden ser problemáticas en la psicopatía; y 3) en cuanto a la validez, ya que no está restringida por baja confiabilidad interevaluador.

No obstante, también presenta ciertos retos; el primero de ellos es la credibilidad de las respuestas brindadas por el evaluado tal como lo plantean Lilienfeld y Fowler (2006),

quienes se cuestionan el por qué alguien intentaría identificar una condición caracterizada por la mentira y manipulación, pidiéndole directamente al individuo que responda honestamente preguntas sobre su condición. Otro de los retos que se presentan, es lograr que el instrumento evalúe correctamente los rasgos afectivo-interpersonales del Factor 1 del PCL-R. Esto debido a que varios instrumentos autoaplicados no han logrado capturar este constructo de esta forma; tal es el caso del Inventario de Personalidad Multifásico de Minnesota (MMPI), específicamente la escala de Desviación Psicopática (Pd) (McKinley & Hathaway, 1944). Asimismo, el Inventario Psicológico de California (CPI) en su escala de Socialización (So) (Gough, 1960), y el Cuestionario Multidimensionalidad de la Personalidad (MPQ), en el super factor de Restricción (MPQ; Tellegen, 1978/1982).

En resumen, el instrumento psicométrico para la evaluación de las características de personalidad psicopática que se construyó, pretende poseer características de evaluación dimensional, con una aplicabilidad tanto en población forense, como en población clínica y general. Además, se pretende normalizar y validar bajo estándares psicométricos rigurosos, con una población costarricense, lo que aumenta las posibilidades de aplicabilidad en un contexto latinoamericano; todo esto contribuyendo a la investigación nacional sobre el tema.

Antecedentes de investigación

A continuación, se detalla la evolución del constructo de la psicopatía; se abordan algunos trabajos que permiten un mejor entendimiento sobre la aplicación de la evaluación psicométrica del trastorno de personalidad psicopática y contribuyen teóricamente al constructo de la psicopatía como tal.

Cabe resaltar que se dividen en categorías tales como antecedentes internacionales y nacionales, esto permitió constatar que hay una carencia de material científico sobre el constructo de la psicopatía a nivel nacional. Contrariamente a la abundancia de material e investigaciones internacionales alrededor del constructo.

1. Antecedentes de investigación internacionales

1.1. *Psychopathy Checklist Revised (PCL-R)*

Durante las últimas décadas, el PCL-R se ha instaurado como la regla de oro para la evaluación y medición del trastorno de personalidad psicopática (Westen and Weinberger, 2004, en Hare et al., 2018), debido a su amplia aceptación y uso en investigación sobre el constructo. Además, cabe resaltar que el modelo del PCL y más tarde el PCL-R, marca un punto donde el estudio del constructo de la psicopatía se torna más científico, al proveer una métrica común, generalmente aceptada, válida y confiable, para la evaluación del constructo (Hare et al., 2018).

Esta escala está compuesta por 20 *ítems* y fue diseñada para evaluar psicopatía en contextos de investigación, clínicos y forenses. Para completar tal escala se utiliza una entrevista semiestructurada, datos procedentes de expedientes, a la vez que otras fuentes de información colateral que permiten inferir y evaluar rasgos de personalidad, así como conductas relacionadas con la concepción tradicional de psicopatía (Hare, 2010).

Los ítems de esta escala se estructuran en dos factores y cuatro facetas. El primer factor se denomina “Interpersonal/Afectivo” y el segundo factor “Estilo de vida/Antisocial”. Dentro del factor 1 se encuentran las facetas 1 “Interpersonal” y la faceta 2 “Afectivo”; por su parte en el factor 2 se encuentran la faceta 3 “Estilo de vida” y la faceta 4 “Antisocial” (Hare, 2010).

Dicha prueba pretende identificar individuos con características psicopáticas: a mayor puntuación existe una mayor probabilidad de que el individuo evaluado se acerque al prototipo de psicópata. Si bien es cierto existe un debate referente a la psicopatía como un constructo categórico o uno continuo, debe tenerse presente que en ocasiones es útil la utilización de puntos de corte categóricos. El punto de corte sugerido por Hare es de 30 puntos en el PCL-R para clasificar a los sujetos como psicópatas (Hare, 2010).

En cuanto a los criterios de confiabilidad mediante la aplicación estándar cuenta con un alfa de Cronbach de .84 al considerar todas las diversas muestras empleadas, y al basarse exclusivamente en la revisión de expedientes con un alfa de .87 al combinar las diversas muestras (Hare, 2010).

Asimismo, la validez de dicha prueba ha sido ampliamente estudiada. Mediante análisis estadísticos ha sido posible que se valide como un constructo coherente, tanto con sentido teórico como clínico. Se ha encontrado que los *ítems* del PCL y el PCL-R son consistentes con los criterios clínicos tradicionales y actuales acerca de los rasgos de personalidad y las conductas que definen el constructo de psicopatía (Hare, 2010).

Las personas con rasgos psicopáticos generalmente comienzan sus carreras delictivas a una edad más corta que otros delincuentes, al menos según lo reflejan el número de contactos formales con el sistema de justicia criminal; tal y como señalan Hemphill et al. (1998), quienes indican que la edad media en el momento de ser condenado por primera vez era de 18,6 años para los delincuentes con una puntuación de 30 puntos o más en el PCL, 19,5 para los delincuentes con una puntuación entre los 20 y los 30 puntos y de 24,3 años para los delincuentes con una puntuación en el PCL de menos de 20, siendo la diferencia estadísticamente significativa.

No obstante, si bien es cierto el PCL-R no fue construido para predecir la conducta del delincuente, actualmente se encuentra firmemente relacionado con instrumentos cuya finalidad es efectivamente para hacer tales predicciones. De hecho, la validez predictiva del PCL-R es suficientemente buena como para impulsar a algunos investigadores a emplearlo como referencia para evaluar la capacidad predictiva de instrumentos diseñados específicamente para tal propósito (Hare, 2010).

1.2. *Psychopathy Personality Inventory (PPI)*

El *Psychopathy Personality Inventory (PPI)* fue desarrollado por Lilienfeld (1990) para detectar características psicopáticas en población no criminal. Está constituido por 187 ítems de autorreporte en escala Likert de 4 puntos, contiene 8 sub-escalas que permiten evaluar facetas de bajo orden de la psicopatía: Egocentricidad maquiavélica, Potencia social, Intrepidez, Resentimiento, Inconformidad impulsiva, Externalización de la culpa, No planificación sin preocupación e Inmunidad al estrés; y brinda un puntaje global de la misma. Asimismo, posee escalas de validez para detectar posibles estilos de respuesta problemáticos (Lilienfeld y Widows, 2005). Este instrumento posee una buena consistencia interna (alfa de Cronbach entre .90 y .93) y una buena confiabilidad test-retest, de $r = .95$. Además, mostró evidencias de validez convergente y discriminante con medidas de autorreporte de psicopatía y comportamiento antisocial. Lilienfeld y Widows (2005, como se citó en Sellbom, et al.,

2018) crearon una versión revisada del PPI, el PPI-R. Esta versión consiste en 154 ítems, en los que se disminuyó el nivel de lectura promedio, se revisó la redacción para eliminar fraseos idiomáticos o culturales, y se mejoró sus propiedades psicométricas en general.

Existen evidencias que indican una correlación mayor del PPI (y PPI-R) con el factor 1 del PCL-R, por sobre el factor 2 (Poythress et al., 1998, como se citó en Patrick, 2006). A pesar de que el PPI se diseñó con base a 8 subfactores, Benning, et al. (2003, como se citó en Patrick, 2006) demostraron que el PPI conforma una estructura conceptual de dos factores, similar a la del PCL-R. Contrariamente a los factores del PCL-R que tienen una moderada correlación, los factores del PPI son ortogonales, de manera que el PPI no comprende un síndrome clásico de psicopatía marcado por signos y síntomas (Patrick, 2006).

1.3. *Self Report Psychopathy Scale (SRP)*

El instrumento *Self Report Psychopathy Scale (SRP)* es una escala de autorreporte diseñada por Hare y colegas con base en el PCL. La versión revisada, el SPR-II, incorporó la noción de los dos factores, donde el primero evalúa características nucleares interpersonales y afectivas de la psicopatía, y el segundo factor evalúa un estilo de vida antisocial e impulsivo (Patrick, 2006).

El SPR originalmente poseía una buena consistencia interna con un alfa de Cronbach de .80, así como buena correlación con diferentes instrumentos de trastorno antisocial y psicopatía (Hare, 1985, como se citó en Patrick, 2006); esto trajo algunas dudas sobre la validez de constructo y la validez discriminante de un trastorno antisocial generalizado (Patrick, 2006).

La última versión, el SPR 4 (Paulhus et al., 2016), consiste en 64 ítems puntuados en una escala Likert de 1 a 5; estos se agrupan en los factores del modelo del PCL. Esta versión está estandarizada en población tanto general como ofensora, posee validez de constructo aceptable en ambas poblaciones, así como una consistencia interna de .89 a .92 en las muestras estudiadas, y buena puntuación test-retest, de .82. El SRP 4 ha mostrado una estructura de cuatro factores: Manipulación Interpersonal, Afecto Insensible, Estilo de Vida Errático y Tendencias Criminales.

Con esta versión, los autores buscaron 1) capturar la estructura factorial del constructo psicopático, 2) reducir el número de ítems relacionados a ansiedad, 3) mejorar la cobertura de la faceta antisocial, y 4) incrementar la confiabilidad de puntuaciones factoriales (Sellbom, et al., 2018). Con esto, se buscaba alinear los factores del SPR a los del PCL-R (Hare et al., 2018).

1.4. *Externalizing Spectrum Inventory (ESI)*

El Inventario del Espectro Externalizante (ESI) (Krueger et al., 2007; Venables y Patrick, 2012) es un instrumento de autorreporte que permite indexar un amplio rango de características y comportamientos problemáticos correlacionados en el dominio de la deficiencia del control de impulsos. Este instrumento organiza comportamientos y características de bajo orden alrededor de factores como propensión desinhibitoria, insensibilidad-agresión, y abuso de sustancias.

Este instrumento es la operacionalización del Modelo de Espectro Externalizante (Krueger et al., 2002; Krueger, et al., 2007; Krueger et al., 2005), abordado más adelante. Está diseñado para el uso con muestras clínicas y no clínicas. Análisis estructurales de las escalas de rasgos de bajo orden demuestran la presencia de un factor comprensivo, de alto orden, que refleja rasgos de desinhibición y una propensión general a problemas de control de impulsos, además de dos sub-factores que reflejan insensibilidad-agresión y uso excesivo de sustancias.

Venables y Patrick (2012) realizan un estudio de validez del ESI en población privada de libertad, la evidencia concuerda con el planteamiento del Modelo de Espectro Externalizante; además, se muestra una relación positiva entre ESI y la psicopatía, evaluado con el PCL-R y el PPI; las asociaciones fueron más evidentes en el Factor 2, desviación antisocial. Adicionalmente, los resultados proveen evidencia para el sub-factor de insensibilidad-agresión como punto de contacto entre el espectro externalizante (este espectro se abordará más adelante en este trabajo) y características nucleares afectivas-interpersonales de la psicopatía.

1.5. *Triarchic Psychopathy Measure (TriPM)*

El modelo Triárquico fue formulado para reconciliar e integrar diversas teorías históricas y contemporáneas, así como operacionalizaciones del constructo de psicopatía.

Patrick et al. (2009, como se citó en Sellbom, et al., 2018). Los autores organizaron los aspectos comunes y únicos de estos modelos dentro de tres amplios dominios fenotípicos: Audacia (*Boldness*) que incluye intrepidez, dominancia social y baja reactividad al estrés; Mezquindad (*Meanness*) compuesta por insensibilidad, explotación o aprovechamiento y desapego interpersonal; y Desinhibición que incluye impulsividad, pobre auto-regulación y baja tolerancia a la frustración.

El TriPM es un instrumento de 58 ítems de autorreporte, creado a partir del modelo Triárquico. El TriPM enfatiza la evaluación de la psicopatía en términos de facetas distinguibles, en lugar de la psicopatía como un constructo unitario y global. Estas facetas son, Desinhibición, con ítems que apuntan a una dimensión más amplia de propensión a la externalización. Mezquinad, comprende impulsividad e irresponsabilidad, tendencias a quebrantar reglas y una dimensión subsidiaria de insensibilidad. Finalmente, Audacia, tendencias de dominancia y falta de miedo, relacionadas con una dimensión bioconductual.

1.6. Levenson Self-Report Psychopathy Scale (LSRP)

Esta escala fue desarrollada para detectar características psicopáticas vía autorreporte en poblaciones no institucionalizadas. Consiste de 26 ítems puntuados en una escala Likert de 4 puntos; estos ítems conforman dos escalas Primaria y Secundaria, paralelas a los Factores 1 y 2 del PCL-R (Levenson, et al., 1995, como se citó en Sellbom, 2018). Levenson et al. (1995, como se citó en Sellbom, 2018) hipotetizaron que el Factor 1 del PCL-R es indicativo de psicopatía “primaria” (Cleckley, 1988), así como el Factor 2 es indicador de psicología “secundaria”.

Más tarde Brinkley et al. (2008, como se citó en Sellbom, et al., 2018) fallaron al tratar de replicar la estructura de dos factores propuesta por los autores originales de la prueba y replicada en estudios previos (Lynam et al., 1999; Brinkley et al., 2001, como se citó en Sellbom, et al., 2018). Investigaciones más recientes han encontrado que el modelo de tres factores presenta un mejor ajuste para el LSRP, Egocentricidad, Insensibilidad y Antisocialidad (Sellbom, 2011; Christian y Sellbom, 2016; Shou et al., 2016; Somma, et al., 2014, como se citó en Sellbom, et al., 2018); sin embargo, Salekin et al. (2014; como se citó en Sellbom, et al., 2018) obtuvieron una mejor validez convergente y discriminante del modelo de dos factores sobre el de tres.

Existen ciertas problemáticas en cuanto a la validez del constructo de esta escala. Varios estudios, como el de Poythress et al. (2010, como se citó en Sellbom, et al., 2018), que han demostrado que el LSRP está mayormente asociado con medidas de psicopatía secundaria y comportamientos antisociales (Factor 2), en lugar de medidas de aspectos afectivos nucleares y características interpersonales propias del Factor 1 del PCL-R.

Asimismo, se ha identificado que el LSRP no cubre rasgos de intrepidez, inmunidad al estrés, o potencia social, asociados a escalas del PPI (*Fearless Dominance*) y TriPM (audacia); esto podría traer implicaciones para el instrumento como medida de psicopatía Primaria al no incluir rasgos y síntomas importantes del constructo psicopático. Finalmente, a pesar de una buena validez de la escala de psicopatía Secundaria/Antisocial, está saturada con demoralización, por una abundancia de ítems relacionados a ira y frustración (Sellbom, 2018).

1.7. *Elemental Psychopathy Assessment (EPA)*

El EPA, construido por Lynam et al. (2011, como se citó en Sellbom, et al., 2018), consta de 178 ítems de autorreporte basados en el modelo de cinco factores de rasgos de personalidad (FFM por sus siglas en inglés). Estos ítems se organizan en escalas que evalúan variantes maladaptativas de 18 rasgos del modelo de cinco factores, que han sido relacionados con la psicopatía.

El EPA se compone de una puntuación total, cuatro puntuaciones factoriales, 18 puntuaciones de subescala y dos puntuaciones de escala de validez, estas pueden ser interpretadas tanto de manera general como a nivel de subescalas. Los cuatro factores son Antagonismo, relacionado al dominio de Amabilidad (*Agreeableness*) del FFM, Estabilidad Emocional, correspondiente a los rasgos del dominio de Neuroticismo. Escalas que reflejan variantes maladaptativas de la faceta de Responsabilidad (*Conscientiousness* en el FFM), cargan en un tercer factor denominado Desinhibición, en el EPA. El cuarto factor, Narcisismo, se asocia con una variedad de rasgos circunscritos en los dominios de Extraversión, Neuroticismo y Responsabilidad.

1.8. *Comprehensive Assessment of Psychopathic Personality (CAPP)*

Cooke et al. (2012) construyen un modelo para el abordaje del trastorno de personalidad psicopática, planteado desde un enfoque dimensional del constructo. El CAPP, *Comprehensive Assessment of Psychopathic Personality* (Evaluación Comprehensiva de la Personalidad Psicopática) describe el constructo psicopático desde un enfoque léxico, a partir de síntomas redactados en idioma natural y descriptivo de rasgos. Más adelante en este trabajo, elaboramos este modelo en mayor detalle.

Sellbom y Cooke (2019, como se citó en Sellbom et al., 2019) desarrollaron el CAPP Self-Report (CAPP-SR), un instrumento de autorreporte basado en la conceptualización del CAPP, este se compone de 3 ítems por cada uno de los 33 síntomas propuestos según este modelo; de esta manera, se asegura suficiente cobertura de contenido de los aspectos léxicos clave para cada síntoma. Estos ítems se configuran en una escala de 0 a 4.

Investigaciones (Sellbom, et al., 2019) han demostrado una buena validez de criterio, .84 y .73 en muestras universitaria y comunitaria respectivamente. Asimismo, el CAPP-SR tiene una buena correlación con otras medidas de evaluación del trastorno psicopático (CAPP Lexical Rating Scale, TriPM, SRP-4, Expanded LSRP, Elemental Psychopathy Assessment—Short Form and Super Short Form y Youth Psychopathic Traits Inventory—Short Form).

En base a estos resultados, se puede apreciar tanto el gran potencial que tiene el CAPP-SR como medida de autorreporte en psicopatía, como el potencial del CAPP como modelo explicativo para el constructo de la personalidad psicopática.

1.9. *Medición en psicopatía*

En general, a excepción del ESI y el reciente CAPP, estos instrumentos o modelos de evaluación constituyen los estandartes de la evaluación psicométrica del constructo de la psicopatía ya que presentan sustancial evidencia de su validez como modelos explicativos del constructo y/o instrumentos de medición y contribuyen con la validez del constructo de la psicopatía (Patrick, 2006; DeLisi, 2016). El ESI y el modelo del CAPP, por su parte, presentan un modelo innovador y prometedor de conceptualización de la psicopatía que permite su abordaje en un espectro externalizante, como se elaborará más adelante.

Según Salvador et al. (2017), debido a que la mayoría de la literatura sobre psicopatía gira en torno a su medida, se concibe como psicopatía lo que miden los instrumentos, lo que ha conllevado a que los modelos subsecuentes se deriven de esos instrumentos. Estos autores llevaron a cabo un metaanálisis en 39 estudios para evaluar la validez convergente de instrumentos derivados del PCL-R, PPI, SPR y LPSP, a nivel unidimensional y de factores. Se pretendía verificar si los instrumentos medían el mismo constructo, y por ende si es un modelo transversal a todos los instrumentos.

Los resultados muestran una validez convergente positiva, significativa y generalizable para los cuatro instrumentos. Además, se determina al PCL-R como prueba de referencia para la evaluación de psicopatía en población clínica, debido a sus evidencias de validez de constructo, algo que, según nuestra óptica, no es lo más apropiado, debido a limitaciones implicadas propias del modelo del PCL.

Por otro lado, se encuentra mayor evidencia para recomendar el PPI como herramienta de evaluación en población sub-clínica o general. Además, la validez convergente del factor 1 es menor, es decir, conceptualmente los instrumentos difieren en este factor; el PCL-R, SRP y LPSP miden factores afectivos e interpersonales, y el PPI mide potencia social, ausencia de miedo e inmunidad al estrés.

Debido a que, en estos instrumentos, PCL-R, PPI, SPR y LPSP, se da mayor peso a las conductas antisociales y se dejan de lado los rasgos de personalidad, Salvador et al. (2016) concluyen en la necesidad de clarificar los límites del papel desempeñado por los rasgos de personalidad y los comportamientos antisociales en la psicopatía. Las conclusiones de los autores evidencian una necesidad de implementar un instrumento de abordaje dimensional para la medición de la psicopatía donde se aborden mayormente las características de personalidad y los elementos del factor 1.

No obstante, es necesario rescatar la importancia de estos instrumentos en tanto brindan fundamentos teóricos sobre el entendimiento que tenemos del trastorno de psicopatía, aportan información sobre qué aspectos deben ser considerados rasgos psicopáticos.

1.10. Viabilidad del autorreporte en psicopatía

Debido a la naturaleza de los rasgos de personalidad psicopática, se ha cuestionado la utilidad de instrumentos de autorreporte para la evaluación psicométrica del constructo. Parecería paradójico utilizar este tipo de mediciones para evaluar un constructo caracterizado por la deshonestidad; sin embargo, el uso de cuestionarios de autorreporte, a pesar de obstáculos y retos, puede ser más fructífero de lo que parece (Sellbom et al., 2018; Lilienfeld y Fowler, 2006). A continuación, abordamos algunas ventajas y desventajas, planteadas por Sellbom, et al., (2018), sobre el uso de medidas de autorreporte para la medición y evaluación de la psicopatía.

El autorreporte es de gran utilidad para la evaluación de rasgos y estados emocionales subjetivos. Específicamente, en psicopatía, rasgos relativamente ausentes como culpa, empatía, miedo y sentimientos de intimación con otros, pueden ser indicadores de la presencia de características psicopáticas; similarmente, rasgos como ira o alienación, más frecuentes en individuos altos en psicopatía, pueden ser captados por este tipo de instrumentos.

La evaluación de estilos de respuesta como impresión positiva o fingir síntomas, es de particular relevancia en individuos altos en psicopatía. A pesar de que estos estilos de respuesta sesgada pueden afectar negativamente la validez del autorreporte para estos individuos, escalas de validez integradas a los cuestionarios, pueden ayudar a detectar este tipo de patrones de respuesta.

En un aspecto económico, las medidas de autorreporte son más fáciles de completar en tanto es el mismo sujeto quien lo hace, además requieren mínimo entrenamiento para los administradores de la prueba.

De mano al punto anterior, la confiabilidad inter-evaluador no es relevante en este tipo de instrumentos, al no requerir juicios de valor de parte de los evaluadores o entrevistadores. Características nucleares de la psicopatía como la falta de empatía y culpa, requieren inferencia clínica de parte del evaluador, de manera que esta subjetividad en medidas basadas en entrevista limita su validez.

La principal desventaja del uso de medidas de autorreporte para detectar rasgos psicopáticos, es esperable, individuos con características de personalidad psicopática tienden

a ser deshonestos y mentir, mentiras que más allá de buscar algún beneficio (mentira instrumental), puede ser simplemente por el disfrute de hacerlo (mentira patológica). E incluso cuando escalas de validez pueden detectar estas tendencias, estas pueden ser susceptibles a que se entrene a la persona a contestar de manera que altere los resultados para dar una imagen más o menos alterada.

Falta de *insight*, se refiere a la incapacidad o fallo de parte de la persona con características psicopáticas, de darse cuenta de la naturaleza y extensión de sus problemas psicológicos. Esta disfunción o falta de *insight*, en cuanto a la capacidad de percibirse como otros lo perciben, pueden afectar la efectividad de algunos ítems de autorreporte, especialmente aquellos que requieren conocimiento verás sobre el impacto del propio comportamiento sobre otras personas.

No obstante, investigaciones han encontrado una fuerte convergencia entre puntuaciones auto-reportadas y reporte de evaluador en medidas de psicopatía con pocas diferencias entre sí (Miller et al., 2011; Fowler y Lilienfeld, 2007 como se citó en Sellbom, et al., 2018). Estos resultados podrían indicar que el impacto potencial de esta falta de *insight* sobre las medidas de autorreporte en psicopatía, puede ser mínimo; sin embargo, esta conclusión puede estar limitada a medidas de autorreporte que no requieran de juicios metacognitivos sobre las implicaciones de la psicopatía sobre sí mismo y otros.

Otra desventaja que mencionan los autores es la problemática de preguntarle a alguien que reporte la ausencia de una emoción, si nunca la ha experimentado. Cleckley (1988) asemejaba la psicopatía con el daltonismo, en tanto al igual que una persona daltónica aprende a referirse a una manzana como roja o una hoja como verde, personas altas en psicopatía pueden aprender etiquetas verbales sobre algunas emociones, aunque erróneamente. Por ejemplo, pueden aprender a referirse como culpa cuando experimentan afecto negativo al recibir un castigo por algún acto antisocial que cometieron, pero lo que están experimentando es arrepentimiento (displacer seguido a ser atrapados), en lugar de remordimiento. Con base en este punto, el reporte de emociones por parte de personas con rasgos psicopáticos altos puede ser impreciso, pero no insincero.

Uno de los principales retos de la construcción de medidas de autorreporte sobre psicopatología en general, es evitar la saturación de desmoralización, una dimensión afectiva de alto orden que refleja una propensión a experimentar emociones negativas de diversos tipos

(Finney, 1985; Tellegen, 1985; Tellegen et al., 2003, como se citó en Sellbom, et al., 2018); a esto podríamos agregar también, la saturación de conductas antisociales. Lo anterior puede provocar una afectación de su validez discriminante para distinguir la psicopatía de otras condiciones marcadas por un comportamiento antisocial.

De acuerdo con Sellbom et al. (2018), alrededor de la evaluación de la psicopatía con medidas de autorreporte, existen ciertas ideas erróneas o concepciones equivocadas, como lo son el requerimiento de respuestas verídicas para garantizar su validez. Sin embargo, Meehl (1945 como se citó en Sellbom, et al., 2018) nota que las respuestas a ítems de auto-reporte pueden o no ser factualmente precisas, pero pueden proveer información de utilidad diagnóstica sobre las apreciaciones de quien responde, sobre sí mismo y el mundo.

Una segunda idea errónea es que personas altas en psicopatía buscan consistentemente crear una impresión positiva en medidas de auto-reporte. Sin embargo, diversas investigaciones han mostrado una tendencia de este tipo de medidas a correlacionar negativamente, de forma leve o moderada, con índices de deseabilidad social y un manejo de impresión positiva (Hare, 1982; Lilienfeld y Andrews, 1996; Marion et al., 2013; Ray et al., 2013, como se citó en Sellbom, et al., 2018). Asimismo, personas altas en psicopatía tienen una concepción diferente de lo que se puede considerar socialmente deseable, pueden percibir como “normal” conductas antisociales que el resto de las personas ven como indeseables (Sellbom, et al., 2018). De la mano con este punto, medidas de estilo de respuesta no dependen enteramente de la varianza de los rasgos, por lo que puntuaciones extremas son probablemente heterogéneas ya que reflejan una varianza relevante a personalidad y psicopatología, aunque una disimulación consciente de parte de algunos respondientes (Sellbom, et al., 2018).

Una última idea errónea es la de que las personas altas en psicopatía son habilidosas para manipular sus respuestas en medidas de autorreporte, por lo que sus respuestas serían incluso menos confiables que las de disimuladores no psicopáticos. No obstante, a pesar de que individuos altos en psicopatía pueden verse más inclinados a sobre reportar o minimizar sus respuestas, según más les convenga (Rogers, et al., 2002, como se citó en Sellbom, et al., 2018), no hay evidencia de que realmente sean habilidosos haciéndolo.

Adicionalmente a estas ideas erróneas, el uso de instrumentos de autorreporte para la medición de psicopatía es afectado por tres problemas empíricos importantes (Sellbom, et al., 2018). Primeramente, cuestionarios diseñados para medir psicopatía no poseen una fuerte

correlación, lo que sugiere que evalúan únicamente aspectos superpuestos del mismo constructo, y por ende estas medidas comparables de psicopatía no son medidas intercambiables.

Finalmente, y como una de las limitaciones más criticadas de las medidas de autorreporte para psicopatía, es que parecen ser medidas de desviación conductual, es decir, conductas antisociales y criminales principalmente. En lugar de medidas de características interpersonales y nucleares afectivas propias de la psicopatía, como la culpabilidad, insensibilidad, falta de amor y egocentricidad; existe un mayor énfasis en el Factor 2 que el Factor 1 según el modelo del PCL. Este problema puede estar explicado porque la psicopatía no es un constructo unitario y es poco probable que logre ser captado por una escala de un inventario de auto-reporte general.

Como han puntualizado diferentes investigaciones (Phillips et al., 2014; Sellbom et al., 2005; Sellbom et al., 2012; Sellbom et al., 2016, como se citó en Sellbom, et al., 2018), una combinación de escalas de rasgos interpersonales y afectivos nucleares de la psicopatía, son típicamente necesarios para lograr un alcance total a la medición del constructo psicopático.

Asimismo, como lo plantean Widiger y Crego (2019b), el futuro de la investigación en la medición de la personalidad psicopática debe dismantelar el constructo psicopático tradicional (PCL) para aislar las facetas o componentes particulares que conforman su validez predictiva, en términos de un modelo general de funcionamiento de personalidad y dejando de lado el enfoque en la evaluación de conductas criminales y antisociales.

En base a los puntos anteriores y al caso del CAPP-SR, podemos destacar la relevancia del CAPP como modelo explicativo del constructo psicopático. Como se abordará más adelante, desde un abordaje dimensional de este trastorno con ítems de bajo nivel, hasta el esfuerzo por no estar cargado con ítems desmoralizantes, el CAPP propone un modelo más completo y adecuado para el abordaje, medición y evaluación de las características de personalidad psicopática.

A lo largo de este repaso, hemos podido observar cómo la evaluación de la psicopatía, vía autorreporte, puede brindar información útil sobre la ausencia de rasgos y estados afectivos, así como la presencia de rasgos desadaptativos, además de presentar ventajas

logísticas. En esta línea, el CAPP-SR, basado en este innovador y completo modelo de abordaje, se postula como una alternativa de gran valor para la evaluación y medición del constructo psicopático.

2. Antecedentes nacionales

Es importante destacar en este punto la escasez de material psicométrico de elaboración nacional, o Latinoamericana para el abordaje y estudio de la psicopatía. La mayoría de los instrumentos usados en el abordaje de este trastorno son instrumentos creados y validados con población norteamericana. Algunos de estos han sido adaptados a población Latinoamericana; sin embargo, esto conlleva limitaciones debido a las diferencias culturales y de contexto de la muestra utilizada.

Con base en estos puntos, se circunscribe la importancia de la creación de un instrumento de medición de la psicopatía, y psicopatología en general, fundamentados en poblaciones del contexto Latinoamericano que se ajuste a las condiciones de la población y a las posibles variables que intervengan en la manifestación de la personalidad psicopática.

De forma similar, el estudio de la psicopatía como constructo, en Costa Rica, es limitado. Entre los trabajos más relevantes destaca el de Saborío (2005), donde, mediante el uso del PCL-R para la medición de la psicopatía, aborda el papel de este trastorno en la configuración de diferentes subtipos de violadores, de acuerdo con el *Massachusetts Treatment Center Rapist Typology* (MTC). Los resultados indicaron un 35% de población con criterios de psicopatía, según el PCL-R, y diferencias significativas entre los violadores no psicopáticos y los psicopáticos, según el MTC, donde estos últimos tuvieron mayor motivación por la oportunidad y la ira generalizada que por factores de tipo sexual.

Por otra parte, Gamboa y Pérez (2004) exploran la asociación entre trastornos y desajuste psicológico con la violencia desplegada por mujeres en la comisión de delitos por los cuales se les sentenció a una condena penal. Mediante la aplicación del PCL-R se evidenció la prevalencia del trastorno en la población y que no había diferencias significativas en la manifestación del trastorno entre ofensoras violentas y no violentas. Además, se determinó un mayor grado de ajuste emocional y social de ofensoras con historial delictivo con tendencia violenta sobre las que tienen predominancia de ofensas no violentas. Sin embargo, al hacer la diferenciación en el tipo de violencia empleada, se notaron características

psicológicas menos adaptativas en las ofensoras instrumentalmente violentas sobre las ofensoras reactivas/hostiles.

Asimismo, existen diferentes trabajos finales de graduación en los que se aborda el constructo de psicopatía en diferentes temáticas relacionadas con la psicología forense. Arce y Gutiérrez (2003) trabajaron la psicopatía en ofensores sexuales, una descripción sobre los niveles de psicopatía en diferentes tipos de violadores de mujeres y la forma en que se comete el delito.

Similarmente, Masís (2008) estudia la psicopatía en ofensores sexuales de menores de edad desde una óptica de la psicología forense; igualmente, Pisoni (2006) abordó la psicopatía en población adolescente, como factor de riesgo para cometer ofensas sexuales. A su vez, Mora y Purcallas (2005) abordaron el trastorno psicopático y su relación con los tipos de violencia en casos de homicidio, mientras que Picado (2007) lo hizo con feminicidas. Reyes (2006), por su parte, estudió la experiencia de culpa y su relación con la atribución causal del dolor o perjuicio de los demás, con una muestra de estafadores con y sin rasgos psicopáticos.

De manera más reciente, Orozco y Quesada (2018) construyeron una prueba psicométrica para la evaluación de factores de riesgo asociados con la competencia o idoneidad mental para portar armas de fuego. Este instrumento de autorreporte consta de 24 escalas de constructos relevantes asociados a un alto riesgo de violencia, y una escala de disimulación. Vale la pena mencionar el estudio realizado por Orozco y Quesada, aunque no relacionado directamente con el constructo de psicopatía, es el más reciente dentro de un contexto de estudios psicométricos.

Estos estudios nacionales evidencian la existencia de características de personalidad psicopática en la población privada de libertad en Costa Rica, pero de igual forma se resalta la necesidad de mayor investigación respecto a este constructo en esta población, debido a que la cantidad de estudios es limitada y queda en evidencia que es un tema que no se ha abordado en los últimos años. Además, se destacan características de personalidad importantes en esta población, las cuales son importantes a tomar en cuenta durante la construcción de *ítems* para la prueba que se pretende construir. Esta información puede contribuir a identificar *ítems* pertinentes para la evaluación de la psicopatía en población costarricense.

Marco teórico - conceptual

1. Trastornos de la personalidad: categóricos vs dimensionales

Se ha planteado la hipótesis (Widiger, 1993; Widiger y Costa, 1994, como se citó en Guay, et al., 2007) de que los trastornos de personalidad en general se conceptualizan como diferentes configuraciones de puntuaciones o valores extremos en rasgos de personalidad, competencias afectivas y cognitivas o procesos neurobiológicos que existen en un continuo con el funcionamiento normal.

Sin embargo, a lo largo de la historia de la clasificación de la salud mental, la conceptualización más utilizada ha sido la del modelo categórico, el cual, según Livesley et al. (1994) se caracteriza por ser dicotómico, es decir, un trastorno o síndrome es considerado como presente o ausente en una persona. Esta decisión es tomada según la cantidad de indicadores o rasgos del trastorno en cuestión, que sean encontrados en la persona. Idealmente, se esperaría que estos indicadores provean un amplio rango de manifestaciones del trastorno.

Estos autores destacan que los rasgos dentro de un mismo trastorno no necesariamente van a covariar, pero todos son pertenecientes y válidos como criterios diagnósticos dentro del mismo. Esto cobra especial relevancia, cuando entendemos este sistema dentro del contexto de la psiquiatría que, al tener sus raíces en la medicina, está influenciada por la aplicación diagnóstica de enfermedades, que siempre ha sido dicotómica, y que permite la posibilidad de un mismo diagnóstico a diferentes personas con distintas cantidades de indicadores detectados.

Por otro lado, Livesley et al.(1994) explican que un rasgo puede ser entendido desde la dimensionalidad, como un largo continuo, tal y como se mencionó anteriormente. En este continuo, se pueden ver reflejadas características mal adaptativas a través de toda la población, según Kotov et al. (2017), quienes afirman que las dimensiones reflejan diferencias en severidad y no en tipo, como pasa con el modelo categórico.

Ante el creciente estudio de la dimensionalidad en la salud mental en las últimas décadas, se ha evidenciado la necesidad de generar un nuevo modelo del estudio de la psicopatología.

Es por ello que Kotov et al. (2017) proponen el modelo *Hierarchical Taxonomy of Psychopathology* (HiTOP), donde se aborda la psicopatología de manera dimensional mediante estudios estructurales. El HiTop, según Mullins – Sweatt et al (2021), es un nuevo sistema diagnóstico, que sintetiza el conocimiento que tenemos de la organización de la psicopatología. Este modelo presenta un sistema de organización jerárquico, de lo más angosto a lo más amplio, que organiza las diferentes dimensiones, el cual se establece de la siguiente manera:

1. Componentes homogéneos: los cuales se definen como constelaciones de síntomas altamente correlacionadas.
2. Rasgos desadaptativos: los cuales son características específicas de personalidad patológicas.
3. Síndromes: los cuales se entienden como un conjunto de componentes o rasgos. Kotov et al. (2017) lo ejemplifican con la ansiedad social, la cual incluye los rasgos de ansiedad por desempeño y ansiedad por interacción. En este apartado también se encuentran la mayoría de trastornos propuestos en el DSM.
4. Subfactores: los cuales son un conjunto de síndromes altamente relacionados.
5. Espectro: que es definida como una gran constelación de subfactores.
6. Súper espectro: la cual se entiende como una conglomeración muy amplia que incluye varios espectros.

El HiTop, según Mullins – Sweatt et al (2021), propone que, en lo más amplio, se encuentran súper espectros, como es el caso de la externalización, por ejemplo, como representación de una conglomeración de aspectos que atraviesan dimensionalmente la psicopatología. Seguidamente, como espectros están presentes dominios amplios de internalización, trastornos del pensamiento, somatomorfo, y externalizante subdividido en desinhibición externalizante y antagonismo externalizante. A su vez, el espectro de internalización se divide en subfactores de miedo y angustia (Kotov et al., 2017).

Asimismo, el Instituto Nacional de Salud Mental (NIMH, por sus siglas en inglés) (s.f.) plantea el modelo *Research Domain Criteria* (RdoC), que también propone un abordaje dimensional de los constructos psicológicos. Estos se agrupan en dominios del comportamiento y funcionamiento humano de más alto rango; que a su vez reflejan conocimiento actual sobre sistemas de emoción, cognición, motivación, y comportamiento

social para, al igual que el modelo HiTOP, proveer un abordaje más integral de los constructos psicológicos.

Parte de la evidencia que se ha encontrado para respaldar la existencia de una dimensión internalizante del espectro, es que el mismo comprende trastornos de ansiedad, depresivos, de estrés post-traumático y alimenticios, así como disfunciones sexuales y trastorno obsesivo compulsivo. Adicionalmente, el súper espectro externalizante comprende trastornos de uso de sustancias, oposicionista desafiante, de conducta, intermitente explosivo, comportamiento antisocial y trastorno de déficit atencional con hiperactividad. El objetivo de este modelo es construir la nosología empíricamente con componentes homogéneos, rasgos desadaptativos y síndromes dimensionales; es necesario hacer la salvedad que se nombran los trastornos con propósitos de comprensión y comunicación, pero todos son concebidos como rasgos y características dentro de la dimensión de la psicopatología.

Paralelamente al estudio del HiTop, se ha estado investigando la estructura de la personalidad patológica, en ella han emergido cinco dominios que parecen integrarla, estos son el afecto negativo, el desapego, la desinhibición, el antagonismo y el psicoticismo (Kotov et al., 2017). La dimensión de afecto negativo es definida por los trastornos de personalidad de dependencia, evitación y de personalidad limítrofe, donde dicha dimensión es un factor común. Mientras que, en el antagonismo, se puede encontrar reflejado en el trastorno de personalidad antisocial, narcisista, histriónico, de personalidad limítrofe y paranoide. En cuanto al desapego, podemos encontrar trastornos de personalidad esquizoide, esquizotípico, y de evitación. Por su parte la desinhibición se ve definida por aspectos como la impulsividad, conductas no premeditadas, e incluso agresión, con lo cual se ha visto correlacionada con el trastorno obsesivo compulsivo, abuso de sustancias y conductas antisociales (Kotov et al., 2017). Por último, el psicoticismo ha sido relacionado por las investigaciones con los trastornos del pensamiento.

En conclusión, debido a que la dimensionalidad permite un mejor y más profundo entendimiento de los constructos, y ante los posibles beneficios que podrían obtenerse al analizar un trastorno como la psicopatía bajo dicho espectro, se ha decidido convertir la dimensionalidad, en uno de los ejes primordiales a la hora de crear el instrumento que se propone en esta investigación, ya que, de esta manera, puede proveer un grado de análisis más específico, en comparación a otras pruebas ya existentes.

2. Críticas al Trastorno Antisocial de la Personalidad y DSM-5

El DSM-5 es el principal referente en salud mental a nivel global, su historia nos lleva hasta los años 1840, según la Asociación Americana de Psiquiatría (APA por sus siglas en inglés), quienes afirman que durante esa década se dio el primer paso hacia la construcción del DSM, con el “censo de locura” de los Estados Unidos. Para el censo del año 1880, ya existían siete categorías de salud mental, sin embargo, no fue hasta el año 1952, que se hizo la primera publicación oficial del DSM, misma que fue impulsada e influenciada por la época post segunda guerra mundial, donde la Milicia de Estados Unidos y la Administración de Veteranos jugaron un rol fundamental en su creación (Psychiatry.Org - DSM History, 2022).

El DSM tiene como principal objetivo estandarizar un lenguaje, análisis e interpretación de síntomas de las enfermedades mentales entre usuarios y países (Carragher et al., 2015) sin embargo a lo largo de las décadas, se ha presentado una serie de críticas por parte de la comunidad científica de la salud mental, y con el pasar de los años, han quedado evidencias sus diferentes limitaciones.

Una de las principales críticas hacia el DSM a lo largo del tiempo, es que este no ofrece mayor información sobre los procesos subyacentes a los trastornos mentales, lo que sí ofrece en su lugar es una descripción detallada de los síntomas que estos conllevan, lo cual hace parecer al DSM a algo más cercano a un diccionario descriptivo que a un manual diagnóstico de psicopatología (Echeburúa et al., 2014).

Debido a esto, durante el pasar de los años ha emergido una creciente frustración entre psiquiatras y psicólogos clínicos, debido a que el DSM desde su perspectiva categórica, muestra un exceso de rigidez y especificidad en sus diagnósticos. Esto en conjunto con la gran cantidad de información clínica relevante que suele perderse al adherirse a la etiqueta clínica, ha conllevado a una constante equivocación a la hora de diagnosticar a los pacientes. Ésta se da desde dos vertientes, una de ellas es el uso del *trastorno de la personalidad no especificado*, el cual implica que los síntomas no encajan con ningún otro trastorno definido. Mientras que en la mayoría de los casos se hacen diagnósticos con una alta comorbilidad entre trastornos, es decir donde coexisten más de un trastorno presente, esto sucede al no poder encontrar explicaciones en un trastorno específico. Este exceso de diagnósticos múltiples ha llegado a enmascarar la complejidad que existe detrás de la psicopatología, lo cual solo refuerza la importancia de pasar a un modelo dimensional, que parece ajustarse y entender más esta

realidad, eliminando así la rigidez que existe actualmente entre los límites de la normalidad y la psicopatología (Carragher et al., 2015).

Ante estas problemáticas, y la evidente necesidad de avanzar en otra dirección, se generó una alta expectativa de que habrían cambios en la nueva versión del DSM, es decir, en la publicación del DSM-V, sin embargo, esto no pudo estar más alejado de la realidad. La nueva versión generó una gran decepción dentro de la comunidad científica, ya que el DSM-V mantiene las mismas definiciones nosológicas de su versión previa (Esbec y Echeburúa, 2015).

Además, hemos visto a través de los años, una creciente patologización de la normalidad (Caponi, 2018), tal y como evidencia Sandín (2013), al mostrar que el DSM en su primera edición describe 106 trastornos mentales, mientras que la actual identifica 216, lo cual es alarmante cuando tomamos en consideración lo deficiente que ha demostrado ser su metodología y las bases de su creación. Con respecto a este último punto, es importante recordar que el DSM tiene un diseño basado en el consenso médico, no en la evidencia empírica (Echeburúa et al., 2014), lo cual es muy grave, ya que esto indica que la comunidad de profesionales de la salud mental está basando sus diagnósticos en trastornos que carecen, o podrían tener un mayor grado de respaldo científico.

El DSM-V sigue siendo politético, lo que significa que al diagnosticar se presenta una lista de síntomas comunes, pero no se requiere que estén presentes todos para confirmar la presencia del trastorno, y ante esto, sigue siendo claramente categórico. Esto hace que exista heterogeneidad dentro grupos diagnosticados con el mismo trastorno, ya que no toma en cuenta las diferencias individuales de las personas, tales como la edad, conjunto de síntomas presentes, o el momento en el que se empiezan a presentar estos (Carragher et al., 2015).

Por si esto fuera poco, en los últimos años se ha generado una fuerte discusión alrededor de acusaciones de conflicto de intereses, ya que se cree que las nuevas versiones del DSM han sido creadas o ajustadas para calzar y potenciar el mercado farmacéutico, ya que con el aumento en el número de trastornos, y la facilidad de los diagnósticos de comorbilidad, se facilita que haya una creciente necesidad en la población de acudir a la medicación, en muchos casos, necesitando más de un fármaco, según su diagnóstico (Esbec y Echeburúa, 2015; Caponi, 2018).

Todo lo anterior evidencia la importancia y la urgencia de avanzar hacia un modelo dimensional en el diagnóstico de la salud mental, inclusive los autores Esbec y Echeburúa (2015, p.179) enlistan una serie de ventajas que este modelo podría aportar:

- 1) Es un modelo coherente con la observación de los límites entre psicopatología y normalidad, lo que permitiría avanzar hacia la eliminación de criterios ambiguos que existen actualmente.
- 2) La dimensionalidad se muestra con una perspectiva mucho más acorde a la complejidad que los trastornos presentan en la clínica.
- 3) La medición dimensional de los trastornos, puede incluir dentro de su metodología aspectos categóricos, pero no es posible hacer lo mismo a lo inverso, ya que la categorización es incompatible con la naturaleza de la metodología dimensional.
- 4) Con respecto al punto anterior, es importante enfatizar que la categorización desde la dimensionalidad ofrece más flexibilidad en los puntos de cohorte, ya que este toma en consideración las diferencias individuales de las personas.
- 5) A su vez, este modelo permite un análisis más pormenorizado de las características de los pacientes, ya que la dimensionalidad no solo evalúa si un síntoma está presente o no, también evalúa la severidad de este, siendo más coherente y acertado en el diagnóstico final.
- 6) Al haber diagnósticos más concretos y acertados a la realidad del paciente, se facilita a los terapeutas el trabajo de intervención en áreas concretas.
- 7) Naturalmente al haber una mayor apertura diagnóstica, en vez de verse forzados a encasillar los síntomas en categorías específicas, se vería una mejora en la fiabilidad Inter jueces.
- 8) Por último, y un punto de gran importancia, es que la comorbilidad podría dejar de ser un problema, esto debido a que el análisis del paciente se ejecuta en función de la combinación de sus rasgos y no de las categorías, tal y como se mencionó en el argumento anterior.

Debido a que existe esta controversia sobre el necesario cambio en el modelo de evaluación de la salud mental desde hace décadas, con el pasar de los años se han logrado ver algunos avances. El paso de la psicopatología categórica a la dimensional ha tenido un avance tímido y lento, pero inevitable. El DSM no puede seguir ignorando los avances en investigación de la psicopatología dimensional, y, pese a que el DSM-V en su sección II *Trastornos de personalidad*, mantiene las categorías diagnósticas heredadas del DSM III y el DSM IV TR, en su apartado III, mismo que está dedicado a los trastornos que requieren más estudio, introduce el modelo alternativo para los trastornos de personalidad (Esbec y Echeburúa, 2015).

Dentro de esta modesta evolución que ha ido presentado el DSM de lo categórico a lo dimensional, podemos encontrar que en los trastornos de personalidad se mantienen las categorías diagnósticas, pero además se incluye una serie de criterios de evaluación distintos:

- Criterio A: que implica una evaluación de alteraciones del funcionamiento de identidad y autodirección, además del funcionamiento interpersonal que incluye la empatía e intimidad (Widiger et al., 2019).
- Criterio B: en este se incluyen 25 rasgos de personalidad maladaptativos, mismos que se organizan dentro del afecto negativo, desapego, antagonismo, desinhibición y psicoticismo (Widiger y Crego, 2019a). Su intención es valorar rasgos desde la normalidad hacia la psicopatología extrema y el nivel de afectación al funcionamiento de la personalidad (Widiger et al., 2019).
- Por otro lado, los criterios faltantes se relacionan con la inflexibilidad (criterio C) y estabilidad a través del tiempo de los trastornos (criterio D), el desorden de personalidad evaluado no se debe a otro trastorno mental (criterio E), el efecto de uso de sustancias psicoactivas o condición médica (criterio F), ni tampoco se debe a un momento del desarrollo de la persona o a una normalidad en un contexto sociocultural específico (criterio G) (Krueger y Hobbs, 2020).
- Para ayudar a los clínicos en la conceptualización y evaluación del criterio A, el DSM V brinda la escala de niveles de funcionamiento de la personalidad (LPFS) la cual evalúa los aspectos de funcionamiento de la personalidad (identidad y autodirección) y funcionamiento interpersonal (empatía e intimidad) en una escala de 5 puntos, de poco o ninguna alteración, hasta alteración extrema (Krueger y Hobbs, 2020). En el criterio B, se aborda el estilo de la personalidad, se conceptualiza a través de 5

dominios de la personalidad, de los que se derivan 25 rasgos o facetas de personalidad. Está se operacionaliza y evalúan mediante el Inventario de personalidad del DSM (Krueger y Hobbs, 2020).

Si bien se puede evaluar de manera positiva el tímido intento de cambio mostrado en el DSM-V para acercarse más a la dimensionalidad, sus falencias y críticas van más allá del paradigma por el cual se rige, también sus descripciones de trastornos han sido altamente criticadas, tal como es el caso del trastorno de personalidad antisocial (TPA), mismo que en teoría vendría a explicar conceptos como el de la psicopatía.

La historia del DSM y su formulación del TPA es muy antigua y extensa, pero siempre ha estado rodeada de cuestionamientos y críticas, tal y como nos lo hacen ver los autores Crego y Widigier (2014), quienes hacen un repaso de su evolución y sus controversias. En el DSM-I se incluyó un “Trastorno Sociopático de la Personalidad”, no hay mucha claridad sobre cuánta influencia tuvo la lista diagnóstica de Cleckley (1988) (Anexo 1), sin embargo, se considera que existe una gran coincidencia sobre los criterios descritos por ambos.

Para el DSM-II se aumentó la definición del “Trastorno Sociopático de Personalidad”, e inclusive se puede argumentar que se acercó más a lo propuesto por Cleckley, ya que se agregan aspectos a la descripción, tales como que estas personas son egoístas, irresponsables, impulsivas, poco capaces de sentir culpabilidad y de aprender de experiencias y castigos (APA, 1968, como se citó en Crego y Widigier, 2014). En el DSM-III se incluyó una lista de criterios de 9 ítems desarrollada por Feighner, el aspecto positivo más importante a destacar en este cambio fue que se empezaron a agregar criterios específicos, como, por ejemplo, para calificar en “pobre historial marital” se requería al menos dos divorcios (Robins, 1966, como se citó en Crego y Widiger, 2014). Incluso Lykken (2007) afirma que no es hasta el DSM-III y su posterior propuesta en el DSM-IV, que se empiezan a alcanzar medidas de consistencia diagnóstica, ya que empezaron a darse los primeros pasos hacía criterios diagnósticos empíricos, y no inferenciales, dado que anteriormente se basaban principalmente en consensos de comités psiquiátricos. Sin embargo, más allá de estos indicios de avance hacia razonamientos más sólidos, seguían faltando criterios de la psicopatía importantes como lo son la falta de culpa, y una mayor claridad en rasgos como mentiroso patológico.

Algo importante a destacar en esta etapa, es que mientras se desarrollaba el DSM-III, simultáneamente estaba siendo desarrollado el PCL, la coexistencia de ambos criterios ha

generado a lo largo de los años una gran cantidad de críticas de un modelo hacia el otro, mismo que se mantiene hasta la actualidad. Originalmente, se señaló que el PCL evaluaba rasgos de personalidad, mientras que el TPA del DSM-III se enfocaba en una lista de actos antisociales, según Hare (1986, como se citó en Crego y Widiger, 2014). Ante ello se podía considerar al TPA como un modelo más cerrado al estar sobre enfocado en la conducta criminal.

Para la publicación del DSM-IV, ya había sido publicada la versión revisada del PCL. El PCL-R se estructuró con dos factores específicos tal y como explicamos anteriormente, siendo el primero enfocado en características de personalidad, y el segundo en el comportamiento antisocial (Hare, 2003). Al analizar ambas propuestas, el TPA del DMS-IV se veía altamente relacionado con el factor dos del PCL-R, lo que le trajo críticas del mismo Hare, quien expresó esto implicaba que las investigaciones que se basaran en el modelo del DSM, iban a perder mucho del componente del aspecto de la personalidad, mientras que en el PCL-R se abarcaban ambos aspectos (Hare, 2003, como se citó en Crego y Widiger, 2014).

El DSM-V no incluye el trastorno de personalidad psicopática; no obstante, sí concibe el TPA (301.7 (F60.2)), igualmente que, en su versión anterior, bajo el Grupo B de los Trastornos de personalidad, el cual se define de acuerdo con los siguientes criterios diagnósticos:

Patrón dominante de inatención y vulneración de los derechos de los demás, que se produce desde los 15 años de edad, y que se manifiesta por tres (o más) de los hechos siguientes:

1. Incumplimiento de las normas sociales respecto a los comportamientos legales, que se manifiesta por actuaciones repetidas que son motivo de detención.
2. Engaño, que se manifiesta por mentiras repetidas, utilización de alias o estafa para provecho o placer personal.
3. Impulsividad o fracaso para planear con antelación.
4. Irritabilidad y agresividad, que se manifiesta por peleas o agresiones físicas repetidas.
5. Desatención imprudente de la seguridad propia o de los demás.
6. Irresponsabilidad constante, que se manifiesta por la incapacidad repetida de mantener un comportamiento laboral coherente o cumplir con las obligaciones económicas.
7. Ausencia de remordimiento, que se manifiesta con indiferencia o racionalización del hecho de haber herido, maltratado o robado a alguien.

2. El individuo tiene como mínimo 18 años.

3. Existen evidencias de la presencia de un trastorno de la conducta con inicio antes de los 15 años. D. El comportamiento antisocial no se produce exclusivamente en el curso de la esquizofrenia o de un trastorno bipolar.

La descripción del DSM-V del TPA es la disconformidad con la conducta legal y ética, egocentrismo, crueldad, falta de preocupación por los demás a través del engaño, irresponsabilidad y temeridad. Además, se incluyó una “variante psicopática” dentro de los criterios, que incluye baja ansiedad, ausencia de retirada y búsqueda de sensaciones, mismos que se relacionan altamente con una personalidad inmune al estrés y desvergonzado estilo interpersonal, lo que se alinea con lo propuesto en algunos espectros para desórdenes mentales en el HiTOP (Wygant et al., 2020). Los dominios con los que se relaciona son el antagonismo y la desinhibición (Esbec y Echeburúa, 2015). A la hora de ser publicada esta versión del TPA, su historial empírico era bastante extenso, sin embargo, las investigaciones parecían haberse decantado mayoritariamente por el constructo de la psicopatía, esto según un estudio de Blashfield y Intoccia (2000, como se citó en Crego y Widigier, 2014), que mostraba que, si bien existe mucha literatura del TPA a lo largo de los años, en las últimas tres décadas, su crecimiento se ha visto estancado.

Adicionalmente, se ha visto una resistencia por parte de los autores del TPA en el DSM-V de apoyarse en el trabajo empírico de Cleckley y de Hare, ya que, si bien el TPA ha buscado acercarse a los modelos de psicopatía propuesto por estos autores, lo han hecho apoyándose en otros autores e investigaciones, principalmente, en la propuesta de Medida Triárquica de la Psicopatía (TriPM por sus siglas en inglés) de Patrick (Crego y Widiger, 2014).

Esto podría deberse a que a pesar de los cambios que se ha hecho el DSM al TPA con el pasar de los años, este parece seguir siendo limitado al compararse con el alcance de la evaluación del PCL-R. Esto queda evidenciado cuando tomamos en cuenta que la mayoría de los casos diagnosticados de psicopatía en prisiones alcanzan los criterios diagnósticos del TPA pautados tanto por el DSM-IV como por el DSM-V, mientras que solo la mitad de los casos diagnosticados con TPA alcanzan los criterios necesarios para ser diagnosticados con psicopatía, esto según Widiger y Crego (2019b).

Dicho esto, es importante destacar que este no es el único trastorno propuesto por el DSM que tiene una alta comorbilidad con la psicopatía, ya que también existe el Trastorno de Personalidad Narcisista, con el que se ha visto una alta correlación (Lynam, 2011, como se citó en Widiger y Crego, 2019b). Inclusive, Hart y Hare (1998, como se citó en Widiger y Crego, 2019b) mencionan que existe un parecido entre la psicopatía propuesta en el PCL-R y el Trastorno de Personalidad Narcisista, sin embargo, ellos consideran que su modelo de psicopatía muestra una mayor severidad, y que es más completo al mostrar dos facetas, de las cuales, solo una muestra alta comorbilidad con el narcisismo propuesto en el DSM.

La realidad es que se ha mostrado que cuando el PCL-R se utiliza para evaluar psicopatía, el Trastorno de Personalidad Narcisista muestra un alto índice de comorbilidad con ambos de sus factores, sin importar la escala que se utilice. Esto cambia, sin embargo, cuando la escala utilizada para evaluar la psicopatía es el PPI-R de Lilienfeld (Widiger y Crego, 2019b), ya que, en este caso, solo correlaciona bien con el dominio de Impulsividad Antisocial del PPI-R.

Widiger y Crego (2019b) explican que la alta comorbilidad del Trastorno de Personalidad Narcisista con la psicopatía tiene que ver con la inclusión de rasgos directamente implicados con el diagnóstico de la misma, como lo son la arrogancia, falta de empatía, incapacidad o negación de reconocer sentimientos y necesidades de otras personas, manipulación, engaño y conductas antisociales. No obstante, sí hay diferencias marcadas, especialmente en el hecho de que la psicopatía incluye criterios de desinhibición, y “baja consciencia”, mismas que se manifiestan en rasgos como poco confiable, e irresponsable. Mientras que el Trastorno de Personalidad Narcisista, al contrario, muestra criterios de “alta consciencia” manifestados en búsqueda de cumplir objetivos, o búsqueda de excelencia académica.

Otro aspecto interesante para considerar que exponen Widiger y Crego (2019b), es que en su momento se consideró agregar algunos criterios de la psicopatía al TPA del DSM-IV, estos eran encanto superficial, arrogancia y falta de empatía, sin embargo, se descartó su inclusión, debido a que se consideró que esto aumentaría la comorbilidad entre el TPA y el Trastorno de Personalidad Narcisista, debido a que estos criterios son fundamentales para dicho trastorno.

Esto es otro indicador más de que a lo largo del DSM ha habido una resistencia a aceptar el modelo de la psicopatía, mismo que a través de los años se ha mostrado más sólido ante la comunidad científica, ya que ha logrado mostrar una mayor confiabilidad y respaldo empírico. Parece ser que el conjunto del TPA y el Trastorno de Personalidad Narcisista podrían alcanzar un nivel de credibilidad bastante alto, pero su conjunto implicaría adoptar un modelo extremadamente similar al de la psicopatía. Lo que es claro, es que mientras siga habiendo este tipo de inconsistencias, los estudios e investigaciones de la comunidad científica se seguirán alineando a la psicopatía antes que al TPA.

3. Super Espectro de la Externalización: la psicopatía dentro de este modelo jerárquico

Las diferencias individuales en la tendencia de contener o expresar impulsos se ha abordado desde la psicología, recientemente, en diferentes constructos desde desinhibición (Clark y Watson, 1999; Gorenstein y Newman, 1980, como se citó en Krueger et al., 2007) hasta impulsividad (Barratt, 1994; Dickman, 1990; Gray, 1981; Whiteside y Lynam, 2001, como se citó en Krueger et al., 2007). Se han observado correlaciones entre estos constructos desinhibitorios de personalidad y trastornos mentales, incluyendo problemas de uso de sustancias y comportamiento antisocial (Sher y Trull, 1994, como se citó en Krueger, et al., 2007).

Según Krueger et al. (2007), el superespectro de externalización y el espectro de internalización reconocen la conexión estructural entre la psicopatología de la infancia y la adulta. Por ejemplo, el comportamiento antisocial temprano se etiqueta como trastorno de conducta, mientras que el comportamiento antisocial grave que persiste en la adultez se denomina trastorno antisocial de la personalidad. De esta forma, Krueger et al. (2005; 2007) define los trastornos externalizantes como aquellos caracterizados por comportamientos desadaptativos que son dirigidos hacia el ambiente del individuo y que resultan en alguna discapacidad o interferencia con el funcionamiento cotidiano; una propensión a expresar el estrés hacia el exterior.

Krueger et al. (2007) plantean análisis covariados de patrones de co-ocurrencia que revelan un factor, una dimensión latente coherente, que enlaza estos trastornos y los distingue de otros como la ansiedad o los trastornos internalizantes (Kendler et al., 1997; Krueger, 1999; Krueger, Caspi et al., 1998, como se citó en Krueger et al., 2007).

De esta manera, la externalización permite una diferenciación entre los trastornos que tienden a manifestarse en comportamientos, en el caso de la psicopatía, antisociales y desadaptativos, y trastornos donde se internalizan los síntomas o consecuencias. Esto permitiría discernir la psicopatía de otros trastornos o incluso detectar o descartar posibles trastornos comórbidos. Finalmente, este modelo se operacionaliza en el instrumento ESI (Venables y Patrick, 2012) que permite una organización de comportamientos y rasgos problemáticos de bajo orden en factores de alto orden mediante una medida de auto reporte.

Otro punto interesante de abordar es la conexión entre la psicopatía en la niñez y la adultez, debido a que la externalización y la internalización, así como la concepción dimensional de los trastornos, reconocen una conexión estructural entre estos tiempos del desarrollo (Krueger et al., 2007). Este aspecto se escapa de los alcances del presente estudio, pero se aprecia la contribución de la concepción externalizante/internalizante de los trastornos para el estudio y conceptualización del constructo de la psicopatía.

El superespectro externalizante del comportamiento está formado por los espectros de externalización por desinhibición y externalización antagonista. Según Krueger et al. (2005) se ha logrado recabar suficiente evidencia para determinar que este se encuentra separado del espectro de internalización, y de los trastornos del pensamiento en los adultos. Este superespectro externalizante, puede ser observado a través de diferentes trastornos propuestos por el DSM-V, como lo son el TAP, el Trastorno por Déficit de Atención e Hiperactividad, Trastorno de Abuso de Sustancias, Trastornos de la Conducta, y el Trastorno de Personalidad Oposicionista Desafiante. Y si bien la psicopatía no está incluida dentro del DSM-V como un trastorno, es evidente que este superespectro también se ve reflejado en sus criterios.

En el caso del espectro de externalización antagónica, se ha encontrado evidencia de que este está ligado a varios rasgos desadaptativos que se reflejan principalmente en relaciones problemáticas con otras personas, así como tendencias a aprovecharse de otras personas sin considerar sus derechos o emociones. Muchos de estos rasgos se ven presentes en la psicopatía, tales como lo son la manipulación, engaño, falta de empatía, sentido de grandiosidad, conductas agresivas, sentido de dominancia, y sospecha. De igual forma, el espectro de externalización desinhibida también ha sido relacionado con múltiples rasgos desadaptativos, pero estos se reflejan principalmente en aspectos de desorganización, conductas antisociales y consumo de sustancias. Algunos de estos rasgos, también están

presentes en la psicopatía, como lo son la impulsividad, irresponsabilidad, y el ser propenso a tomar riesgos (Kotov et al., 2017; Krueger et al., 2005; Krueger et al., 2021).

Es claro que al analizar el superespectro de externalización en la psicopatía, es más fácil encontrar mayor evidencia en el Factor 2 de comportamiento antisocial, según el cual se puede argumentar que la psicopatía tiene una tendencia externalizante, que se ve evidenciada y se manifiesta en estos comportamientos. Es por ello por lo que la dimensionalidad de la externalización es un concepto crucial que será abordado como un eje primordial en esta tesis.

Como se apreció anteriormente, la psicopatía es un constructo muy amplio y complejo, e incluso recientemente han surgido nuevas conceptualizaciones de este. De esta manera, es de suma importancia el abordaje del constructo desde una visión que compile todos estos aportes para garantizar una apreciación más completa del mismo.

4. Un modelo dimensional de la psicopatía: CAPP

El uso de diferentes medidas basadas en diferentes conceptualizaciones del trastorno, o medidas muy diferentes a pesar de ser concebidas bajo la misma conceptualización, pueden generar dificultades para unificar resultados. De la misma manera que el uso de una sola medida basada en la misma conceptualización puede dejar poca claridad en cuanto a sus hallazgos, si reflejan propiedades de la medida o características subyacentes del constructo.

De esta manera, Cooke et al. (2012) plantean la necesidad de un modelo conceptual para el trastorno de personalidad psicopática; un modelo que capture las mayores características de este, de manera que funcione como guía para entender, evaluar y traducir entre las diferentes conceptualizaciones y operacionalizaciones del constructo.

Los autores se basan en seis supuestos clave sobre los trastornos de personalidad para la definición y construcción de un constructo de trastorno de personalidad psicopática (Cooke et al., 2012).

Los síntomas de un trastorno de personalidad deben reflejar una desviación personal, no una desviación social o cultural. Modelos más tradicionales como el del PCL o el DSM, poseen un gran énfasis en síntomas conductuales expresados en conductas antisociales y criminales. Sin embargo, con el fin de no reducir el constructo y su medición a poblaciones específicas (población clínica o privada de libertad) donde la expresión de estas conductas es

más común, el modelo conceptual debe definirse en términos de rasgos patológicos de personalidad, en lugar de este tipo de conductas.

Los síntomas del trastorno de personalidad deben definirse en términos atomistas. Similarmente al punto anterior, estos deben reflejar características de personalidad a un nivel básico, perturbaciones funcionales de la personalidad; en lugar de características más complejas y de alto nivel.

Los síntomas del trastorno de personalidad deben definirse en términos del idioma natural. Se sigue la hipótesis léxica de que el grado en el cual características de personalidad y atributos particulares son representados en el lenguaje, está ligado a su importancia para transacciones dentro de ese lenguaje. Es decir, las diferencias importantes en transacciones humanas, como los son trastornos y síntomas de la personalidad, son codificadas en idioma natural.

Los síntomas del trastorno de personalidad deben ser definidos en términos que reflejen su naturaleza dinámica, en el sentido que la personalidad patológica fluctúa a través del tiempo (Tyrer, 2005, como se citó en Cooke et al., 2012). Los rasgos de personalidad que definen un constructo deben ser sensibles a esas fluctuaciones con el fin de mantenerse vigentes a lo largo del tiempo.

Los síntomas del trastorno de personalidad pueden estar organizados jerárquicamente. Síntomas de bajo nivel pueden tener importantes asociaciones teóricas o empíricas que les permitan agruparse de formas significativas.

Los síntomas del trastorno de personalidad deben ser comprensivos. Una definición conceptual debe incluir síntomas primarios y excluir los secundarios, en tanto son los primeros los que circunscriben la naturaleza del constructo. Ante la dificultad de determinar si un síntoma es primario o secundario, es mejor ser sobre-inclusivo.

Con base en estos puntos clave, se evidencia la necesidad de la re-definición del constructo psicopático. Una definición construida teniendo en cuenta tanto los aportes como limitaciones de modelos anteriores. Ante esta necesidad, Cooke et al. (2012) plantean un modelo explicativo, dimensional, para el constructo de la psicopatía, Comprehensive

Assessment of Psychopathic Personality (CAPP); Evaluación Comprehensiva de la Personalidad Psicopática.

El CAPP se construyó en cuatro etapas (Cooke y Logan, 2018), primeramente, se realizó una extensa revisión de literatura para identificar los síntomas primarios, aquellos que se consideran centrales o prototípicos de la psicopatía, excluyendo secundarios. Subsecuentemente, se realizaron entrevistas a expertos en el tema para asegurar que no quedaran síntomas por fuera de los ya identificados.

Este modelo propone el abordaje de la psicopatía de manera dimensional, basado en estos síntomas característicos del trastorno de personalidad psicopática recopilados de diversos modelos explicativos sobre el constructo. Síntomas que fueron deconstruidos y definidos en términos de un lenguaje natural y descriptivo de rasgos, en un proceso de triangulación con tres sinónimos o frases adjetivales. Contrariamente, en comparación con las definiciones más tradicionales, como las del PCL, donde cada definición de ítems y síntomas es extensa y se superponen entre sí en términos de contenido.

Esta aproximación léxica provee una metodología para identificar dimensiones de variabilidad en personalidad, distinciones capturadas en las sutilezas y matices de esos adjetivos descriptores de rasgos. De la mano con este punto y desde una premisa de que es importante en el desarrollo de modelos, el ser sobre inclusivo, se incluyeron todos los síntomas potencialmente relevantes, a pesar del desacuerdo de algunos expertos (Cooke y Logan, 2018); esto, con el propósito de construir un modelo verdaderamente completo y expresar síntomas a lo largo de toda la dimensión del constructo.

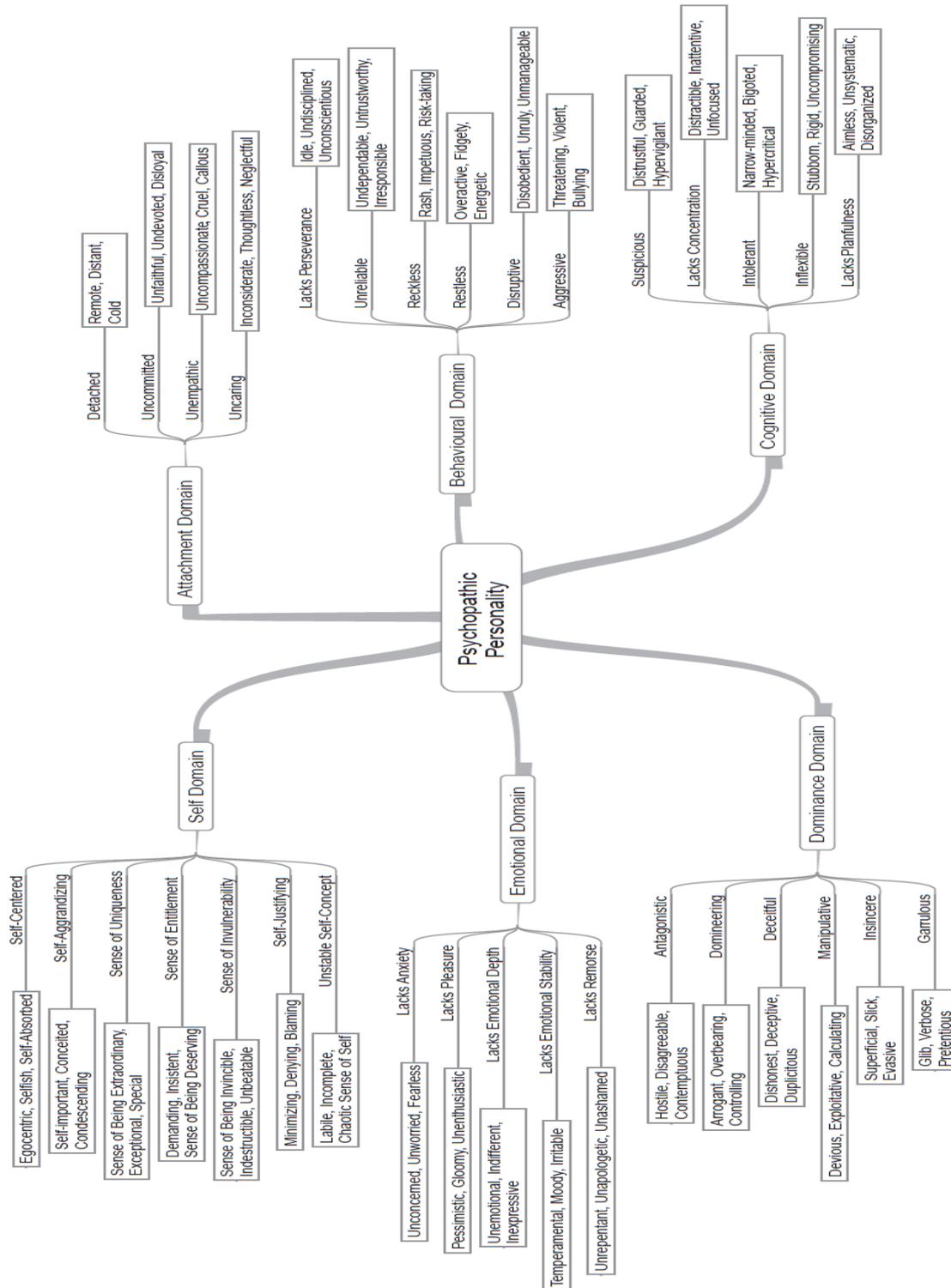
Finalmente, estos 33 síntomas que componen el CAPP se categorizaron, a su vez, en seis dominios de personalidad y funcionamiento psicológico. A pesar de que la evidencia y literatura alrededor del modelo del PCL sugiere que los síntomas psicopáticos reflejan tres dominios de funcionamiento psicológico (interpersonal, afectivo y conductual) (Hare et al., 2018), una evaluación de contenido de los síntomas del CAPP sugiere la posibilidad de una conceptualización más extendida y refinada de los dominios de la psicopatología.

La figura 1 presenta el modelo completo del CAPP, los 33 síntomas del trastorno de personalidad psicopática, definidos por tres adjetivos o frases adjetivales, agrupados en los

seis dominios funcionales: Apego, Sí Mismo, Emocional, Conductual, Dominancia y Cognitivo.

Figura 1

Comprehensive Assessment of Psychopathic Personality (CAPP) (Cooke et al., 2012).



El dominio de *Apego* se refiere a los vínculos emocionales con otras personas; refleja dificultades de afiliación interpersonal, con un enfoque en la intimidad y aceptación por otros que se busca alcanzar y mantener en la interacción interpersonal. En este dominio se circunscriben los siguientes síntomas: desapegado, no adquiere compromisos, sin empatía e insolidario.

Sí Mismo se refiere a la conciencia de la propia identidad, la percepción que tiene el individuo de sus propios rasgos, esquemas, habilidades, cualidades y necesidades, a la vez de como éstas influyen en los roles sociales y la forma de relacionarse con otros. Este dominio refleja problemas con este sentido de identidad e individualidad, y por consiguiente en aspectos como planteamiento de metas y la interacción con otras personas. Este dominio incluye los síntomas Centrado en sí mismo, Autoagrandecimiento, Sentido de unicidad, Sentido de derecho, Sentido de Invulnerabilidad, Autojustificación y Autoconcepto inestable.

El dominio *Emocional* abarca problemas de regulación emocional, con un enfoque en el tono, profundidad y qué tan apropiada es la respuesta emocional. En este dominio se incluyen los síntomas de Falta de ansiedad, Sin capacidad para experimentar placer, Falta de profundidad emocional, Falta de estabilidad emocional y Falta de remordimiento.

Dominancia es un dominio que refiere a conductas interpersonales, métodos por los cuales el individuo compromete a otros para conseguir sus propias necesidades. En este ámbito, refleja dificultades con esta agencia interpersonal, se enfoca en el grado de poder o control que la persona intenta tomar en esos intercambios interpersonales. Antagónico, Dominante, Engañoso, Manipulador, Insincero y Charlatán son los síntomas de este dominio.

El dominio *Conductual* se enfoca en la regulación de conductas, incluyendo el fallo en establecer estrategias adaptativas para lidiar con hasta las tareas más cotidianas de forma sistemática, consistente y planeada. Refleja problemas que puede tener la persona en la organización de actividades dirigidas a una meta. Los síntomas de este dominio son Inconstante, Poco fiable, Imprudente, Intranquilo, Disruptivo y Agresivo.

Finalmente, el dominio *Cognitivo* se centra en acciones y procesos mentales, cómo el individuo centra y dirige su atención, codifica y procesa información, organiza pensamientos y hace atribuciones. Refleja problemas de flexibilidad y adaptabilidad mental en síntomas como Falta de concentración, Suspica, Intolerante, Inflexible y Falta de planeamiento.

En resumen, el CAPP fue desarrollado para delinear los contornos del trastorno de personalidad psicopática. Mediante un acercamiento “bottom-up”, integra perspectivas de diferentes fuentes empíricas y teóricas, generando un modelo de disfunción de personalidad específico a psicopatía. Contrariamente a un acercamiento “top-down”, en el que la definición de un constructo depende de los insights, experiencias y visión de su postulante, el CAPP permite una clarificación de la naturaleza y los límites del constructo, al ser concebido desde una revisión de la literatura, la opinión de expertos, los análisis de estos hallazgos y la observación de casos (Cooke y Logan, 2018).

Tal y como mencionan Widiger y Crego (2019b), una tarea para futuras investigaciones es la de aislar facetas o componentes del constructo psicopático que componen su validez predictiva. Este esfuerzo puede lograrse en función de modelos más generales del funcionamiento de personalidad, o mediante el uso de instrumentos de evaluación que no se basen en un historial criminal y que provean subescalas de los diferentes componentes de la psicopatía. Asimismo, sería valioso investigar muestras con alto o moderado nivel de psicopatía con una ausencia de comportamiento criminal.

El CAPP se propone como una forma de abordaje y evaluación del constructo de psicopatía de forma dimensional, en donde su definición se hace en términos de extremos funcionales de esta dimensión; dando así, espacio para una forma de evaluación no basada en los comportamientos antisociales o criminales.

De esta manera, este modelo propuesto por el CAPP puede ser de gran utilidad para indagar cómo combinaciones específicas de rasgos psicopáticos pueden resultar en diferentes grados de severidad a lo largo de la dimensión del constructo psicopático, y más específicamente, comportamientos y conductas antisociales o criminales.

Problema de investigación

Como se ha indicado anteriormente, el estudio psicométrico de trastornos psicológicos es escaso en Costa Rica, más aún el estudio del trastorno de personalidad psicopática. Esto implica la carencia de material de evaluación para este trastorno en poblaciones forense y general. De mano a esto, también se carece, a nivel nacional, de material metodológico y antecedentes para la construcción de instrumentos psicométricos.

Existen varias adaptaciones latinoamericanas de instrumentos, como el PCL-R que se ha adaptado en Chile (Abalos et al., 2004) y México (Ostrosky-Solís et al., 2008), sin embargo, estas muchas veces no se ajustan a las particularidades de la población, o son medidas basadas en la entrevista y revisión de expedientes, lo que dificulta su aplicación. Un instrumento de auto reporte puede aligerar cargas en el costo de aplicación o para situaciones en las que no se cuenta con un expediente exhaustivo del individuo.

Asimismo, los modelos más recientes que proponen un abordaje dimensional de los trastornos de la personalidad los plantean en un continuo, de manera que un instrumento basado en este abordaje permitirá una indagación más eficaz en términos de la evaluación de la funcionalidad de la persona. En contextos forenses esto contribuye a responder preguntas legales asociadas a la funcionalidad de los individuos en la sociedad.

De esta manera, se lleva a cabo la construcción de un instrumento de auto reporte basado en un modelo dimensional de la personalidad, que cumpla con las necesidades lingüísticas de la población costarricense, desde un abordaje con técnicas cualitativas y cuantitativas. Se espera que esto permita brindar datos normativos sobre la población forense a nivel nacional, además de estadísticas de prevalencia del trastorno en la población.

1. Objetivos generales

- Desarrollar un instrumento de medición psicológica de auto reporte que evalúe características de personalidad psicopática desde una perspectiva dimensional en población costarricense.
- Proporcionar evidencias de validez y confiabilidad sobre el instrumento psicométrico que se diseñe, en una muestra de población general y forense costarricense, así como datos normativos que permitan su aplicación en la misma.

2. Objetivos específicos

- Operacionalizar un constructo de psicopatía que tome en cuenta aspectos dimensionales del mismo que capture su naturaleza multifactorial y sus distintas manifestaciones conductuales, afectivas e interpersonales a lo largo del espectro en población general y forense.

- Brindar evidencias de validez de constructo en distintos tipos de población (forense y general).
- Proveer información sobre evidencias de confiabilidad de la prueba en la evaluación de rasgos de psicopatía en población general y forense costarricense.
- Desarrollar un instrumento de evaluación del constructo de la psicopatía que cuente con normas de interpretación obtenidas a partir de población costarricense y que sirva para ser utilizado en distintos contextos de evaluación en nuestro país.

3. Objetivo externo

- Proporcionar un manual de uso técnico para aplicar un instrumento de evaluación psicológica sobre características de personalidad psicopática de tipo dimensional en población general y forense costarricense.

Diseño metodológico

A continuación, abordaremos en detalle los aspectos metodológicos y logísticos que se llevaron a cabo y posibilitaron este trabajo.

1. Tipo de estudio

Las características no observables de los fenómenos psicológicos son definidos como constructos teóricos, conectados por medio de teorías con las conductas observables y otros constructos; de esta forma, las relaciones de los constructos latentes con hechos observables permiten la cuantificación y establecimiento de relaciones empíricas entre ellos (Martínez-Arias et al., 2014). Estos mismos autores indican:

para la medida de estos constructos latentes se requiere algún instrumento que ponga de relieve los niveles variables del constructo y de algún modelo que relacione lo que se observa en el mundo real (las conductas observables) con el atributo o constructo que se pretende evaluar, que existe como parte de una teoría (p 20).

De esta forma, el presente trabajo se circunscribe bajo la noción de un estudio psicométrico, en tanto realiza un abordaje de un constructo psicológico, como lo es el trastorno de personalidad psicopática, mediante la indagación a través del auto reporte de

características interpersonales, afectivas y de estilo de vida antisocial, según lo planteado por Hare (2010), Krueger et al. (2007), Lilienfeld (1990), con el uso de un instrumento de medición.

2. Procedimiento general para la construcción de pruebas y su validación

El proceso de construcción de la prueba se compuso de las diferentes fases planteadas de acuerdo con las necesidades propias del instrumento que se construyó; no obstante, se tuvo como fundamento las fases de construcción de *tests* planteadas por Martínez-Arias et al. (2014), así como por el proceso propuesto por la Asociación de Investigación Educativa de América (AERA) y la Asociación Estadounidense de Psicología (APA) (2014); esto debido a que estos representan guías claras y completas, las cuáles proveen insumos provechosos para la construcción del instrumento propuesto.

Estas fases son: 1) especificaciones del estudio, 2) descripción de la población y muestra, 3) sobre construcción, validación y pilotaje, 4) diseño y creación de *ítems*, 5) selección y revisión de *ítems*, 6) ensamble de la prueba, 7) desarrollo del protocolo de aplicación, 8) estudio piloto, 9) evidencias de validez de criterio, 10) estudio de campo y aplicación de la prueba, 11) análisis de datos, 12) desarrollo del protocolo de uso y datos normativos, 13) confidencialidad y protección a los participantes, y 14) la retribución a la población. A continuación, se detallan cada una de estas fases del proceso.

3. Especificaciones del estudio

3.1. Identificación del propósito y finalidad del test

La prueba que se construyó tiene como finalidad la evaluación de características del trastorno de psicopatía en población forense y general adulta, con el fin de ser utilizada para evaluaciones de riesgo de violencia en situaciones forenses, evaluación de idoneidad laboral en distintos de puestos, evaluaciones para portación y manejo de armas de fuego, entre otros.

Las puntuaciones de una prueba construida con estas características permitirían una evaluación dimensional de características del trastorno de personalidad psicopático para tomar decisiones de distinta naturaleza (diagnóstico diferencial, clasificación o tamizaje). A

partir de estas puntuaciones, se contaría con insumos para tomar decisiones legales, judiciales, laborales o clínicas respecto a la persona evaluada.

3.2. Especificidades para la aplicación

Como primera especificación del instrumento diseñado se encuentra su característica de ser de modalidad de autorreporte. En cuanto a la población, se restringió la aplicación a personas, mujeres y hombres, adultas, con un nivel de lectura de sexto año de escolaridad primaria y que posean un adecuado nivel de lectura, para que permita la facilidad de comprensión de los reactivos de la prueba, o si su grado de escolaridad es igual o menor a secundaria incompleta.

Es importante resaltar que, desde la construcción de ítems, se ha basado el proceso desde una conceptualización dimensional del constructo psicopático, como el planteado por el CAPP (Cooke, et al., 2012; Cooke y Logan, 2018), en donde este se concibe como un continuo cuyos extremos representan puntos de funcionalidad menos o más adaptativa.

4. Descripción de la población y muestra

De acuerdo con Prieto y Muñiz (2000), un tamaño de muestra en un estudio se considera grande al ser igual o mayor a 500 sujetos, de esta manera, inicialmente se planteó trabajar con una muestra de 400 personas de población forense y 400 personas de población general.

Debido a limitaciones abordadas más adelante (apartado: Limitaciones), se realizó la aplicación de la prueba construida a 18 personas privadas de libertad y 432 personas de una muestra general, buscando obtener resultados totales que se acercaran lo más posible a la suma sugerida por estos autores, con la finalidad de conseguir análisis estadísticos confiables.

El instrumento se validó con una muestra de población general del país que sea económicamente activa y reside mayormente en el Gran Área Metropolitana (GAM). Entre los criterios de inclusión se encuentra el tener un rango de edad que vaya de los 18 años en adelante y comprensión de lectura acorde por lo menos al sexto año de escolaridad primaria, según su grado de escolaridad.

Para la recolección de la muestra se utilizaron herramientas de cuestionarios virtuales, tales como “Google Survey” y posteriormente se utilizó de forma definitiva “Limesurvey”, además se generó un formulario por Excel que fue distribuido en casos específicos en los cuales era de mayor facilidad de uso para las personas que iban a llenar el test. El test fue compartido por medio de redes sociales, tales como “Facebook” e “Instagram”, tanto de forma general, como en grupos específicos universitarios.

Además de estas publicaciones, se hizo una solicitud a la Escuela de Psicología de la UCR para que hicieran pública la investigación y compartieran mediante sus redes el test para que la población general pudiera llenarlo. En conjunto con este esfuerzo, se hicieron acercamientos a diferentes profesores de la carrera de psicología, de igual forma en la UCR, para que estos incluyeran el test como parte de su planeamiento del curso, aprovechando de esta forma la necesidad de llenar la prueba como una plataforma para explicar temáticas del curso, y a su vez, colaborando con nuestra investigación.

Finalmente, obtener datos de población forense fue considerado como algo crucial para esta investigación, si bien hubo dificultades de alta complejidad para obtener estos datos, mismas que serán explicadas en las limitaciones, se logró conseguir colaboración por parte de psicólogos forenses del Centro de Atención Institucional Jorge Arturo Montero Castro, para que el test fuera aplicado a personas privadas de libertad, específicamente del ámbito de convivencia F.

De esta forma, y por estas mismas limitaciones, el tipo de muestreo utilizado fue no probabilístico. Se aceptó la participación de toda persona dispuesta a colaborar con la investigación, siempre y cuando cumpliera con los requisitos mencionados anteriormente.

5. Confidencialidad y protección a los participantes

Las personas que participaron en las diferentes etapas de este proyecto completaron un consentimiento informado con el cual autorizan su participación en el mismo. Además, se aclara en este, que su participación es voluntaria, que no les perjudicó ni benefició de ninguna forma y que la información recabada será utilizada con propósitos meramente de investigación; además se resalta el derecho a discontinuar su participación en cualquier momento. La información recabada será mantenida de manera confidencial y será resguardada de manera segura (ver consentimientos informados en anexos 4 y 5).

Cabe resaltar, para la población forense, el anonimato de las respuestas brindadas; además, las evaluaciones realizadas no van a incidir de ninguna forma en algún proceso legal o judicial. Asimismo, el no participar o decidir interrumpir su participación no conlleva ninguna consecuencia de ningún tipo.

6. Retribución a la población

A manera de retribución a la población general, población privada de libertad y al sistema penal costarricense, se compartirán los resultados de la investigación, describiendo las características de la población y los principales resultados del estudio.

7. Sobre construcción, validación y pilotaje

La fase de construcción de la prueba, en las entrevistas cognitivas, y la fase de pilotaje, detalladas más adelante, se realizaron con población general mediante un muestreo intencional, debido a que es una etapa previa a la validación. En estas fases, a excepción de las entrevistas cognitivas, se incluyeron 60 personas de ambos sexos, desde los 18 años y de diferentes niveles educativos.

Las entrevistas cognitivas se realizaron a 4 personas, de la muestra general mediante una selección intencional, estas personas se seleccionaron debido a su nivel de experticia y la retroalimentación que podían brindar en el marco de construcción de la prueba. Entre ellas, se realizó la entrevista a una profesional en lingüística y un profesional en psicología, para lograr obtener un mayor entendimiento de las interpretaciones que podrían estar teniendo los ítems, y a su vez, obtener retroalimentación de posibles cambios y sugerencias que podrían implicar una mejora a la hora de plantear obtener información a través de los ítems del test.

Posteriormente, se realizó un nuevo corte al alcanzar las 100 personas (57 mujeres y 43 hombres), con el fin de elaborar una guía preliminar de interpretación. Para la elaboración de esta guía, se desarrollaron datos normativos con base en estos 100 datos. La guía de interpretación se basa en puntajes T, que indican elevaciones moderadas y severas, lo que se correlaciona con las definiciones de cada escala y permite la interpretación de los puntajes.

Tabla 1

Muestra respectiva para cada fase metodológica.

Fase	Muestra
Entrevista cognitiva	Población general (N: 4)
Etapa de pilotaje de la prueba ensamblada	Población general (N: 60) Nota: Para los casos en los que este pilotaje mostró propiedades psicométricas apropiadas, serán consideradas parte de la muestra de la etapa “Aplicación de la prueba”.
Desarrollo de guía de interpretación	Población general (N: 100)
Aplicación de la prueba	Población general (N: 432) Población Forense (N:18)
Estudio de validez de criterio	Población general: Se les aplicó el MMPI-2 (N: 59 - 60), el IDPP (N: 70-71), PAI (N: 58), NEO PI-R (N: 59 - 60) y el UPPS-P (N: 140-141).

Fuente: Elaboración propia

8. Diseño y creación de ítems

Se creó una serie de ítems basados en los conceptos definidos teóricamente, principalmente bajo el modelo de psicopatía propuesto por Cooke et al., (2012) con la concepción del CAPP; para ello, fue necesario entrar en contacto con los autores para solicitar formalmente que se compartiera el modelo teórico traducido y validado en español por el investigador Gerardo Flórez. Dicha solicitud fue aprobada, esto fue fundamental para que los ítems fueran generados con mayor rigurosidad científica.

Sin embargo, fue de indudable importancia rescatar las contribuciones de Hare (2003) en el PCL-R, el de Lilienfeld en el PPI (1990) y el de Krueger et al. en el ESI (2007). De igual

manera, estos instrumentos se tomaron como fundamento para el diseño y construcción de los ítems, debido a la evidencia que sustenta su pertinencia para el estudio del constructo de psicopatía.

El proceso de construcción de *ítems* se realizó con base en los 33 síntomas propuestos por el CAPP (Cooke et al., 2012). Fundamentados en la definición de cada síntoma (adjetivos y frases adjetivales), se construyeron reactivos alrededor de la misma, preguntas orientadas a cuestionar sobre cada síntoma en específico. En cuanto al formato de los *ítems* y las reglas de puntuación, toda la prueba se construyó con ítems de escala tipo Likert con una puntuación de cada uno en escala de frecuencia de 1 a 4 (Nunca, Pocas veces, Muchas veces, Casi siempre).

Cada ítem se categorizó en pensamiento, experiencia, conducta o actitud. De esta manera se tiene claridad qué tipo de evocación se hace con la pregunta. Si bien, esta categoría estuvo determinada por la pregunta en sí, se trató de mantener un equilibrio sobre el tipo de evocación e incluir al menos una pregunta de cada categoría para cada adjetivo o frase adjetival de cada síntoma.

A su vez, estas preguntas se plantearon pensando en toda la dimensionalidad del constructo psicopático y el síntoma específico sobre el que se pregunta. De esta forma, durante la construcción se categorizaron las preguntas en niveles de severidad alto, medio y bajo, con el fin de cubrir el síntoma en toda su dimensión. Con esto se pretendió evitar sesgos de criminalización de los ítems, así como asegurar el abordaje de rasgos psicopáticos en el extremo más funcional de la dimensión y en la población general.

Cabe destacar, también, que, con el fin de crear reactivos eficaces y alineados a los objetivos de este trabajo final de graduación, el proceso de construcción de ítems partió de los propuestos de Cooke et al. (2012) para la definición de los trastornos de personalidad (mencionados anteriormente en el apartado: Un modelo dimensional de la psicopatía: CAPP).

9. Selección y revisión de ítems

Una vez construidos los ítems, se realizó una revisión de redacción para asegurarse que fueran consistentes con los síntomas que abordan, así como asegurarse que fueran gramaticalmente correctos y entendibles.

Adicionalmente, los *ítems* seleccionados fueron puestos en evaluación, utilizando un grupo representativo, el cual debe tener características en común con la muestra, como, por ejemplo, rangos de edad concordantes, capacidad de comprensión de lectura y niveles de escolaridad similares. El primero de los procesos a los cuales fueron sometidos los *ítems*, es la entrevista cognitiva.

Es laborioso reportar conductas y actitudes relevantes en cuestionarios estandarizados, que no pueden tomar mucho tiempo; además, el contexto de la aplicación y el cuestionario en sí pueden dar información a los sujetos irrelevante para la investigación, distraerlos, limitar sus estrategias para recuperar la información solicitada o incluso influir en sus respuestas. En este sentido, Smith y Molina (2011) argumentan la importancia de conocer más a fondo lo que acontece mentalmente durante el proceso de responder a un reactivo para poder incorporar este conocimiento en la evaluación de estos.

La entrevista cognitiva es una herramienta de investigación que se basa en la noción de que para contestar un cuestionario primero se debe comprender el significado que se transmite en este; después se inicia el proceso para recuperar la información de la memoria autobiográfica, luego se recurren a diferentes estrategias para estimar la respuesta adecuada y, así, se emite la respuesta (Smith y Molina, 2011). Este modelo de Tourangeau (1984, en Smith y Molina, 2011), plantea las fases de procesos interrelacionados de a) comprensión, b) recuperación, c) estimación, y, d) ejecución de la respuesta.

De esta manera se realizaron entrevistas cognitivas a cuatro participantes de la población general, por conveniencia e igualmente tratando de estratificar esta sub-muestra, siguiendo los lineamientos planteados por Smith y Molina (2011). Estas personas brindaron retroalimentación valiosa para el escrutinio y revisión de los *ítems* desde sus campos respectivos (psicología y filología). Se inicia con la consigna con la cual se instruye a los participantes (ver anexo 3), se indica que es una entrevista semiestructurada individual en un ambiente controlado en la que se solicita completar el cuestionario y se interroga sobre aspectos que intervienen o interfieren en el proceso de contestar las preguntas.

En el proceso de entrevista, se utilizaron las dos técnicas básicas de la entrevista cognitiva mediante un sondeo concurrente, solicitar a los entrevistados decir en voz alta todo lo que piensan mientras completan el cuestionario, y posteriormente, contestar preguntas sobre el cuestionario. Este segundo punto, mediante pruebas cognitivas del reporte verbal que

fueron aplicadas a los aspectos centrales del cuestionario: introducción, los reactivos y las opciones de respuesta para cada reactivo. El objetivo fue detectar problemas en las fases del proceso pregunta-respuesta (comprensión, recuperación, estimación y ejecución de la respuesta) (Willis, 2005, como se citó en Smith y Molina, 2011). Se pueden usar pruebas de a) parafraseo, b) juicios de confianza, c) recuerdo, d) especificación y e) sondeo del proceso mental, para detectar problemas en el procesamiento del cuestionario.

10. Ensamble de la prueba

A raíz de los múltiples ítems generados, se buscaron y unieron aquellos que fueron representativos de las variables que se buscan evaluar. Se buscó construir diferentes escalas, con la finalidad de cubrir todas las variables contempladas, y garantizar confiabilidad y consistencia interna en la prueba. Estas debieron cumplir criterios de contenido y psicométricos. Posteriormente, se modificaron aquellos *ítems* que fueran demasiado similares a otros, o bien, que pudieran sugerir las respuestas de algunas de las otras preguntas presentes en el test. Una vez creados los prototipos de escalas, fueron sometidos a pruebas piloto.

El proceso de revisión del instrumento se llevó a cabo, al presentar los *ítems* a un experto en la materia, un psicólogo forense, mediante lo cual se recibió retroalimentación, para garantizar que el material no cayera en infrarrepresentación del constructo, o bien, que no se esté insertando variables irrelevantes al test. De igual forma, desde una perspectiva en psicometría, con el fin de que estos evalúen la propuesta metodológica, y con ello se realizaron posteriormente las correcciones necesarias para hacer la prueba óptima y funcional.

Adicionalmente, se llevó a cabo una aleatorización de los ítems, esto con el fin de que los mismos estuvieran distribuidos a lo largo de la prueba. Esto se realizó mediante la función de Excel “ALEATORIO.ENTRE”.

11. Desarrollar procedimientos y materiales para administrar y evaluar el test

Para esta fase es primordial tomar en cuenta tanto la población objetivo de la prueba, como la población que sería capaz de suministrar el test. Las instrucciones de aplicación deben ser claras para todo aquel que reciba o imparta la evaluación.

En las instrucciones se especificó que las preguntas serían alrededor de conductas, pensamientos y actitudes propias a lo largo de la vida. Con base en esta premisa, las

instrucciones se redactaron haciendo énfasis en que las personas deben contestar pensando en experiencias ocurridas a lo largo de su vida, no sólo situaciones recientes.

Asimismo, se recalca que no hay respuestas correctas o incorrectas y el interés es más bien conocer de forma honesta la respuesta al enunciado. Se repasan también las opciones de respuesta y una breve descripción de estas, así como indicaciones propias de la plataforma. Finalmente, se menciona nuevamente, la seguridad de la información y respuestas proveídas. A continuación, las instrucciones completas:

Este cuestionario trata sobre conductas, pensamientos y actitudes que las personas adultas pueden presentar en algún momento de su vida.

Lea cada una de las frases del formulario y seleccione con su puntero la opción que se ajusta más a su situación y experiencia personal:

- **NUNCA:** Si lo que dice la frase a lo largo de su vida ha ocurrido **NUNCA**
- **POCAS VECES:** Si lo que dice la frase a lo largo de su vida a ocurrido **POCAS VECES**
- **MUCHAS VECES:** Si lo que dice la frase a lo largo de su vida ha ocurrido **MUCHAS VECES**
- **CASI SIEMPRE:** Si lo que dice la frase a lo largo de su vida ocurre **CASI SIEMPRE**

Solo debe marcar una de las cuatro opciones de respuestas posibles. **NINGÚN** enunciado debe quedar sin responder. **AL RESPONDER, POR FAVOR RECUERDE EXPERIENCIAS QUE HAYAN OCURRIDO A LO LARGO DE SU VIDA Y NO SOLO RECIENTEMENTE.** Por favor, sea lo más sincero posible al responder cada uno de los enunciados y trate de hacerlo de manera precisa.

Si se equivoca y quiere cambiar su respuesta solamente marque con su puntero la nueva opción, el cuestionario no le permitirá elegir más de una opción correcta por enunciado, o continuar sin haber respondido alguno de los ítems. El cuestionario cuenta con varias páginas de ítems, asegúrese de llegar al final de ellas para enviar los resultados. Al final del cuestionario verá un mensaje de agradecimiento y el botón para finalizar, asegúrese de marcarlo.

Antes de iniciar queremos informarle que los datos que se recogerán a continuación serán almacenados utilizando tecnología de encriptación para garantizar su anonimato y serán utilizados con fines estadísticos únicamente. SUS DATOS NUNCA SERÁN TRATADOS DE FORMA INDIVIDUAL.

12. Estudio piloto

La versión inicial del instrumento constaba de 458 ítems, aunque no todos ellos fueron utilizados en la versión final del instrumento; la idea de diseñar y construir más ítems de los que se utilizarán finalmente es poder someterlos, más adelante, a un proceso de escrutinio y selección mediante análisis estadísticos con el modelo *Rasch* (Rojas y Montero, 2012), que será abordado más adelante, eligiendo así, aquellas preguntas que sean capaces de capturar de mejor manera la variable que está siendo evaluada.

La revisión del *test* partió de los procesos de creación y escrutinio de ítems. Sin embargo, una vez generado y probado, el test debió someterse a revisiones, en búsqueda de aspectos por corregir o mejorar. Estos pueden ser de contenido, evaluación o bien, de vocabulario (lo cual puede estar sujeto a aspectos culturales).

Una vez construido y ensamblado el instrumento, se entró una fase de pilotaje en una muestra de la población general para verificar la comprensión de los reactivos de la prueba y revisar otros elementos relacionados a la aplicación, o a la misma prueba que puedan mejorarse.

A esta fase de pilotaje, se le realizaron análisis estadísticos bajo el modelo Rasch, específicamente modelo de ajuste de Rasch interpretado con base en valores de Infit y Outfit, análisis de correlación punto-media, análisis de funcionamiento diferencial y análisis de los componentes principales de los residuos (ACPR) (abordados más adelante en el apartado: Análisis de datos), para evaluar la calidad métrica de los ítems, su validez de contenido y confiabilidad (Rojas y Montero, 2012), con la finalidad de identificar aspectos de los constructos que no sean relevantes para el test, problemas de formato en los ítems, o bien, errores en los criterios de puntuación utilizados. En los casos en los que el pilotaje mostró propiedades psicométricas apropiadas, las aplicaciones fueron consideradas parte de la muestra de la etapa “Aplicación de la prueba”; es decir, se incluyeron a la muestra final del estudio, filtrando solamente los ítems de la versión final del instrumento.

Posteriormente a los análisis iniciales, se redujo la cantidad de ítems a 351, distribuidos entre las 33 escalas. Estos ítems presentaron propiedades psicométricas adecuadas dentro de la muestra del pilotaje, por lo que permanecieron en la versión final del instrumento, versión usada para la aplicación final a la totalidad de la muestra.

13. Estudio de campo y aplicación de la prueba

Una vez finiquitados los detalles y posibles correcciones de la prueba, encontrados en la fase de pilotaje, se procedió a la aplicación de esta a la muestra objetivo. La prueba se aplicó a una muestra de 18 personas privadas de libertad del Centro de Atención Institucional Jorge Arturo Montero Castro, ámbito de convivencia F, y 432 personas de población general, seleccionadas bajo los criterios establecidos anteriormente.

Debido a limitaciones logísticas que serán detalladas más adelante en este trabajo, la aplicación de las pruebas a la población general se realizó en su mayoría mediante la plataforma digital Limesurvey, un hipervínculo con la prueba se compartió con los participantes. Algunos grupos a los que se tuvo acceso (grupos de clases a quienes se les solicitó su participación), lo completaron directamente en un archivo Excel. Ambas modalidades incluían el consentimiento informado y requerían su aprobación para continuar.

Finalmente, la evaluación de las 18 personas privadas de libertad que participaron, fueron evaluadas por psicólogos forenses del mismo centro penitenciario. Esta evaluación se realizó con un cuadernillo de preguntas y hojas de respuesta.

14. Evidencias de validez de criterio

Adicionalmente a la prueba que se pretende construir, se aplicó el Inventario Multifásico de Personalidad de Minnesota MMPI-2, el Inventario Dimensional de Psicopatología de la Personalidad (IDPP), el Personality Assessment Inventory PAI, el Inventario de Personalidad NEO PI-R y la escala de impulsividad UPPS-P, pruebas psicométricas de evaluación de características de psicopatológicas de la personalidad, con el fin de recabar evidencias de validez de criterio. Estas pruebas se aplicaron a personas de la muestra general utilizada para la validación de la prueba que se construye, con el fin de examinar la validez convergente y validez de constructo de la prueba construida. La cantidad

de personas a las que se le aplicó cada prueba varió dependiendo de la disponibilidad de la muestra (ver Tabla 1).

Inicialmente, se pretendía aplicar el PCL-R, PCL:SV y PID-5 a participantes de la muestra forense y general, respectivamente. No obstante, debido a las limitaciones que serán abordadas más adelante (ver apartado: Limitaciones), se optó por utilizar pruebas robustas cuyo objetivo o algunas de sus escalas es detectar rasgos o características que se encuentran dentro de la dimensionalidad que compone el constructo de psicopatía, y que por ende son de interés para determinar la validez de criterio. Estas pruebas serán enlistadas y explicadas a continuación.

14.1. MMPI-2

El MMPI-2 es un instrumento de medición psicométrica altamente utilizado a nivel mundial y con un historial muy longevo, mismo que le ha aportado solidez a lo largo de los años con sus modificaciones, esto teniendo en cuenta que su primera edición fue publicada en el año 1942 (Nichols et al., 2011). Esta prueba psicométrica contiene múltiples escalas que evalúan diferentes psicopatologías y sus diversos síntomas. Adicionalmente, el MMPI-2 es utilizado en ámbitos de selección de personal, clínico y forense. Es por ello por lo que se consideró como una excelente herramienta para un eventual análisis de correlación, específicamente por las siguientes escalas (Nichols et al., 2011; American Psychiatric Association, 2013):

- **Desviación psicopática (Dp):** Propensión a las conductas antisociales, oposición a la autoridad, búsqueda de inmediata gratificación, poco control de impulsos, impaciente, poca tolerancia a la frustración, centrado en sí mismo, exhibicionista, otros reactivos reflejan su falta de interés por la mayoría de las normas sociales y morales de conducta.
- **Hipomanía (Ma):** Ambición exagerada, extraversión, periodos de excesiva actividad, alucinaciones o delirios de grandeza, relaciones superficiales, pocos límites, hostilidad y agresión.
- **Prácticas antisociales (PAS):** Problemas de conducta, irrespeto a las reglas, deshonestidad, prácticas deshonestas o criminales, falta de empatía, desdén.
- **Agresividad (AGR):** Agresividad ofensiva o predatoria y urgencia hostil por dominar, derrotar y destruir a otros.

- Desinhibición (DIS): Amplia dimensión de poco control conductual, poca modulación de impulsos, espontaneidad, flexibilidad cognitiva y moral, poca tolerancia al retraso de la gratificación.
- Comportamiento antisocial (RC4): Uso y abuso de sustancias, comportamiento delictivo, absentismo, relaciones conflictivas y no gratas.
- Activación hipomaníaca (RC9): Agresión hostil, rapidez al enojo, reacciones vindicativas a provocaciones, y anuencia a provocar e intimidar a otros.
- Antisocial: Dificultades en las áreas de identidad, autodirección, empatía, intimidad, así como los rasgos patológicos de antagonismo y desinhibición: manipulación, insensibilidad, engaño, hostilidad, asunción de riesgos, impulsividad e irresponsabilidad.

14.2. NEO PI-R

El NEO PI-R (Costa y McCrae, 2008) es una prueba de medición psicométrica que se ha desarrollado durante las últimas cuatro décadas, siendo publicada su primera edición en 1985. Este instrumento se hizo con el objetivo de volver operativo el modelo de los cinco factores de personalidad, y sus escalas han sido sometidas a arduas investigaciones del ámbito clínico y también con población normativa.

Este instrumento está formado por cinco factores, que a su vez se subdividen cada uno de ellos en otros seis. Para este estudio se tomaron en cuenta todos los factores del NEO PI-R, en miras de obtener información pertinente de la validez discriminante. Estos factores son los siguientes (Costa y McCrae, 2008):

- Neuroticismo: Afectividad negativa, lo cual incluye enojo, hostilidad, ansiedad y volatilidad, además está vinculada con impulsividad y dificultad para tolerar la frustración.
- Extraversión: Sociabilidad, actividad, habilidades de comunicación y afecto.
- Amabilidad: Personas bondadosas, confiadas, serviciales, indulgentes y altruistas.
- Apertura: Personas curiosas, imaginativas, dispuestas a aceptar valores e ideas poco convencionales.
- Responsabilidad: Confiabilidad, disciplina, honestidad, persistencia, seguimiento a figuras de autoridad.

14.3. PAI

El Inventario de Evaluación de Personalidad (PAI), es un instrumento de autorreporte diseñado para ayudar el proceso de diagnóstico, tratamiento, planeamiento y tamizaje de psicopatología, pero es utilizado ampliamente en contexto de justicia criminal (Gardner et al., 2015). Se enfoca en rasgos psicopatológicos relacionados al estrés, rasgos de trastorno de personalidad y necesidades de tratamiento. El PAI consta de 344 ítems en escala Likert de 4 puntos, estos, a su vez, se dividen en 22 escalas completas, compuestas por subescalas; en el presente trabajo se eligieron las escalas que por su relevancia teórica y empírica se relacionan con la psicopatía (Edens, et al., 2000; Edens et al., 2001; Miller et al., 2010; Taylor y Lang, 2006, como se citó en Gatner, et al., 2018; Morey, 2011; Ruiz, et al., 2018).

- Manía (MAN): Síntomas afectivos, cognitivos y conductuales de la manía e hipomanía.
 - MAN-A: Nivel de actividad
 - MAN-G: Grandiosidad
- Rasgos antisociales (ANT): Egocentrismo, empatía limitada, actitudes y conductas antisociales.
 - ANT-A: Conductas antisociales
 - ANT-E: Egocentrismo
 - ANT-B: Búsqueda de sensaciones
- Agresión (AGR): Conductas y actitudes relacionadas a la agresión y hostilidad.
 - AGR-A: Actitud agresiva
 - AGR-V: Agresiones verbales
- Dominancia (DOM): Rasgos controladores, sumisos o autónomos en sus relaciones interpersonales. Independencia, asertividad, liderazgo y capacidad para gestionar relaciones.
- Afabilidad (WRM): Cordialidad, empatía, retraimiento, desconfianza, sociabilidad, solidaridad, afecto y paciencia en relaciones.
- Antagonismo: Tendencia a la hostilidad, rasgos manipuladores, engañosos y grandiosos.
- Desinhibición: Orientación hacia la gratificación inmediata, comportamientos no limitados por consecuencias futuras o aprendizaje previo. Irresponsabilidad, impulsividad y conductas arriesgadas.

14.4. IDPP

El Inventario Dimensional de Psicopatología de la Personalidad (IDPP) es un instrumento autoaplicado para la evaluación de psicopatología (Orozco y Quesada, 2018), en base al modelo de psicopatología propuesto por Krueger (Krueger y Markon, 2014, como se citó en Orozco y Quesada, 2018). Se identificaron 24 dimensiones de psicopatología de la personalidad relevantes para el objetivo de la investigación. Con estas 24 dimensiones se crearon 24 escalas mediante la operacionalización de los constructos asociados a las mismas, las cuales representan tres espectros: Internalización, Externalización y Psicoticismo.

Similarmente a las pruebas anteriores, algunas escalas del IDPP se constituyen como síntomas del constructo psicopático, de manera que se considera valioso correlacionarlas con las escalas del instrumento construido. Estas escalas son: Búsqueda de atención, Manipulación, Engaño, Grandiosidad, Irresponsabilidad, Impulsividad, Búsqueda de sensaciones, Distractibilidad, Potencial de adicción, Insensibilidad, Desinhibición, Antagonismo, Desapego, Afecto negativo y el factor de Psicoticismo con sus respectivas escalas (a excepción de Disimulación) (Orozco y Quesada, 2018).

14.5. UPPS-P

La Escala de Comportamiento Impulsivo UPPS-P (por las siglas en inglés de las escalas que lo componen) (Lynam et al., 2006) se basa en la propuesta de la impulsividad como un constructo multifacético y multidimensional, compuesto por cinco rasgos de personalidad impulsiva. Estas dimensiones/escalas, detalladas a continuación, se dividen en 59 ítems de escala Likert de 4 puntos:

- Urgencia negativa: Tendencia a actuar apresuradamente ante extremas emociones negativas
- Falta de premeditación: Tendencia a actuar sin pensarlo
- Falta de perseverancia: Incapacidad para permanecer concentrado en una tarea
- Búsqueda de sensaciones: Tendencia a buscar experiencias emocionantes y novedosas
- Urgencia positiva: Tendencia a actuar de forma precipitada ante una extrema emocionalidad positiva

Bajo el modelo de impulsividad planteado en el UPPS-P, se plantea una concepción integrativa de la impulsividad, al evaluar estas cinco formas o tipos de impulsividad.

La impulsividad es un constructo que se circunscribe como parte de los síntomas de la dimensión del constructo de psicopatía; tanto el modelo del PCL (Hare, 2010) como el del CAPP (Cooke, et al., 2012) incluyen la evaluación de la impulsividad dentro de sus escalas. De esta manera, se consideró oportuno realizar análisis de correlación de la prueba construida con el UPPS-P, para indagar de qué manera correlacionan sus escalas.

15. Análisis de datos

Se realizaron análisis del ajuste de los ítems a las distintas escalas, así como distintos procedimientos que permitan detectar los ítems que aportan a la validez de contenido y homogeneidad de las escalas. De igual forma, se llevaron a cabo análisis sobre el posible funcionamiento diferencial de los ítems (DIF) en factores como el género. Asimismo, un análisis sobre el funcionamiento de la escala Likert utilizada en las alternativas de respuesta de las escalas. El análisis de los datos se realizó con el programa WINSTEPS versión 5.0.2.0, para el modelo Rasch (Linacre, 2018), y Jamovi 2.3.13 para los análisis restantes.

Se realizaron algunos análisis con Teoría Clásica de los Tests, como los correspondientes a la fiabilidad de las distintas escalas, mediante el cálculo del coeficiente Alfa de Cronbach. También se llevó a cabo, un análisis factorial exploratorio, con la finalidad de obtener datos con respecto a la validez de constructo del instrumento. De igual manera, se realizaron análisis de T de Student, procedimiento que permitirá encontrar diferencias entre los grupos muestrales (Ritchey, 2008), en este caso, por género. Además, se hizo uso del análisis de varianza (ANOVA), con el objetivo de realizar comparaciones entre diferentes grupos de edad y sus rasgos psicopáticos. Finalmente, se utilizó el coeficiente de correlación de Pearson, para obtener datos de validez de criterio convergente y divergente, con respecto a las otras pruebas aplicadas en el proceso. A continuación, se expondrá con más detalle las implicaciones de estos análisis.

15.1. Análisis de modelo Rasch

Se realizó un análisis con el modelo de Rasch, que se enmarca dentro de la Teoría de Respuesta al Ítem (TRI) (Rojas y Montero, 2012), la TRI se basa en el supuesto de que existe

una relación entre los valores de las variables que miden los ítems y la probabilidad de acertarlos. De esta manera, el modelo Rasch se utiliza para medir un fenómeno latente a partir de las puntuaciones de diferentes sujetos a una serie de ítems (Plá, 2009); el modelo establece comparaciones entre las personas y los ítems para determinar si existe un ajuste a un modelo unidimensional con un constructo o variable latente subyacente que se expresa en la respuesta correcta de los ítems. Así, detecciones de desviaciones en el desempeño o valores de las personas o ítems, funciona como un mecanismo de control de calidad para establecer que tan buen ajuste tiene cada ítem con el constructo latente y si los resultados se pueden interpretar como mediciones de intervalo (Saborío, 2022).

15.2. Análisis de ajuste de ítems y personas

Mediante estadísticos basados en los residuos, o la diferencia entre las respuestas observadas y las esperadas, este modelo permite determinar qué tan apropiado es un instrumento para medir un constructo (Cadavid et al., 2007, como se citó en Saborío, 2022). Por una parte, el valor de Information-weighted fit statistics (Infit, por sus siglas en inglés), se interpreta como una media cuadrática ponderada de residuos, sensible a, y que permite detectar patrones de respuesta irregulares, respuestas no esperadas a ítems calibrados cerca del nivel de habilidad de la persona, o ajuste interno (Saborío, 2022). Con base en esto, valores altos en el Infit indican que los ítems tienen un rendimiento inadecuado para las personas que están destinados, lo que amenaza la validez de la medición (Linacre, 2022).

Por otra parte, Outlier-sensitive fit statistics (Outfit, por sus siglas en inglés), similar pero contrariamente al Infit, detecta el ajuste externo en patrones de respuesta inesperados a reactivos lejanos al nivel de habilidad de las personas, en los extremos de la distribución de la muestra (Saborío, 2022). Es una medida basada en la estadística de chi-cuadrado convencional, y sensible a respuestas inesperadas en los reactivos que pueden ser muy fáciles o difíciles para las personas; de forma que un alto valor Outfit puede indicar respuestas aleatorias por parte de los sujetos de bajo rendimiento (Linacre, 2022).

Las medidas principales de Infit y Outfit son los cuadrados medios (MNSQ), donde 1.0 es el nivel de ajuste perfecto entre datos y el modelo, valores altos son una amenaza mayor para la validez que los valores bajos. Valores significativamente menores a 1.0 indican dependencia o determinismo entre los datos observados; mientras que valores significativamente mayores a 1.0 indican una posible aleatoriedad o ruido en los datos.

Linacre (2022), propone un criterio de ajuste de Infit y Outfit dentro del rango 0.5-1.5, para considerar los ítems como productivos. Valores en un rango 1.5-2.0, indican un desajuste moderado pero que no degrada la medida; y valores superiores a 2.0, sí muestran un desajuste severo. Este criterio fue el utilizado para la depuración de ítems dentro del presente trabajo.

El Infit y Outfit también es un criterio aplicable a las personas, donde valores por debajo de 2.0 en más del 85% de las personas, con un valor promedio aceptable de estos indicadores, permite llegar a la conclusión que la mayoría de las personas poseen un ajuste básicamente adecuado al modelo (Bartolomé y Prieto, 2017; Prieto et al., 2018, como se citó en Saborío, 2022). Similarmente, Miguel, et al. (2019), indican que un porcentaje bajo (18.50%) de personas con infit y/o outfit por encima de 2.0, aún se circunscriben en el umbral superior de aceptabilidad psicométrica para el modelo de Rasch.

15.3. Análisis de Funcionamiento Diferencial

El Análisis de Funcionamiento Diferencial (DIF por sus siglas en inglés), es un análisis fundamental para determinar la habilidad de un ítem de ser generalizable en diferentes grupos, aspecto que es un fundamental en la validez de medición (Myers et al., 2006). Este análisis lo que implica es una evaluación del grado en el que el significado de las mediciones, puede ser generalizado entre distintos subgrupos que forman parte de una población.

El DIF puede ser encontrado cuando participantes con el mismo nivel de severidad del atributo medido, no tienen la misma probabilidad de puntuar el ítem hacia cualquiera de sus extremos, es decir leve o severo. Desde el enfoque de Rasch, el DIF se calcula una vez que se ha ajustado las diferencias del grupo, y posteriormente se lleva a cabo una correlación de Bonferroni del nivel de significancia (Linacre, 2022).

La importancia de evaluar el DIF recae en que, cuando se tienen grupos dentro de una muestra, con niveles de severidad similares del atributo medido, pero que contestan de manera altamente distinta, se generan sesgos de puntaje, mismos que podrían llegar a amenazar la unidimensionalidad del constructo. Con lo cual, el DIF determina si estos grupos o subgrupos de la muestra, responden de forma sistemáticamente distinta a alguno de los ítems del test.

Linacre (2022), indica que en el modelo de Rasch, para los test de significancia como lo es el DIF, es importante tomar en cuenta tanto la significancia estadística, como la

diferencia sustantiva entre los grupos comparados. Es por ello que se recomienda que el valor de DIF CONTRAST, es decir la diferencia entre el nivel de dificultad entre los grupos comparados, sea de al menos 0.5 para ser considerado relevante. Mientras que el valor “Prob.”, mismo que indica la probabilidad de observar este contraste debido al azar, debe ser menor a 0.05. Con lo cual, los ítems que se ajustaron a estos rangos fueron eliminados de la prueba.

15.4. Análisis de correlación punto-media

En cuanto a la Correlación Punto-Media, análisis que es similar al de correlación ítem-total del CTT (Miguel et al., 2019), Linacre (2022) sugiere que los puntajes por arriba de 0.4 son un indicador de que cada uno de los ítems contribuye a la definición de un constructo común. Con lo cuál, con el objetivo de obtener una medida fundamentalmente unidimensional, se eliminaron todos los ítems con una correlación con la variable inferior a .30 (Fisher y Prieto, 2017).

15.5. Análisis de Componentes Principales de los Residuos (ACPR)

El Análisis de Componentes Principales de los residuos (ACPR), sirve como una técnica que permite a los investigadores evaluar si los datos obtenidos, se ajustan o no al modelo de Rasch. El objetivo del ACPR es evaluar si los datos cumplen con el supuesto de unidimensionalidad, esto mediante el uso de las medidas de Rasch, para explorar si existe una dimensión adicional en un conjunto de datos, y adicionalmente, entender cuál podría ser la fuerza de esa evidencia (Boone y Staver, 2020).

Entonces, se puede entender la función principal del ACPR como el intentar extraer el factor común que explica en mayor medida la varianza residual, esto bajo la hipótesis de que existe dicho factor (Linacre, 2022). Si se descubre que este factor está detectando ruido aleatorio, entonces se tendría evidencia de que no hay factores significativos adicionales en los datos.

Para alcanzar el supuesto de unidimensionalidad del modelo de Rasch, es necesario que los valores propios (Eigenvalue) de los residuos sean inferiores a 2.0, ya que esto indica que el primer contraste está a nivel del ruido, y, por ende, se afirma que el ruido observado fue aleatorio. Si los valores propios de los residuos son mayores a 2.0, entonces es necesario

revisar los ítems que puedan estar generando una segunda dimensión, esto al observar los diferentes análisis que se están exponiendo en este apartado.

Adicionalmente, según Bartolomé y Prieto (2017), se considera que es suficiente una dimensión cuando la dimensión principal de la misma explica como mínimo un 20% de la varianza (Reckase, 1979).

15.6. Análisis de confiabilidad

El modelo Rasch proporciona índices que permiten determinar si hay suficientes ítems esparcidos a lo largo del continuo, así como distribución o dispersión de las habilidades de las personas en este mismo continuo (Bond y Fox, 2007; como se citó en Saborío, 2022). De esta forma, el índice de fiabilidad indica la replicabilidad del ordenamiento esperado de una muestra si se aplica otro conjunto de ítems para medir el mismo constructo. La fiabilidad de las personas refiere a una jerarquía de habilidad de los participantes, en sus respuestas.

Según los autores, la fiabilidad requiere tanto ítems que capturen diferentes rangos de habilidad de las personas de la muestra, como personas que presenten distintos rangos de habilidad en sus respuestas para que haya una dispersión adecuada en el continuo. En el caso de la fiabilidad, se espera un puntaje lo más cercano a 1 posible, siendo aceptable desde .7, y en el caso de la separación se espera un puntaje por arriba de 2; si se cumplen estos criterios, indica, que la prueba pudo diferenciar entre grupos de dificultad, habilidad de ítems y personas (Linacre, 2022).

Igualmente, el Alfa de Cronbach se considera como uno de los análisis más confiables, y además más utilizados para la validación del constructo de una escala, mismo que cuantifica y analiza la correlación que existe entre los ítems de la misma, según González y Pazmiño (2015). Adicionalmente, estos mismos autores destacan que el consenso en la comunidad científica, indica que un valor de Alfa de Cronbach entre 0.70 y 0.90, es un buen indicador de consistencia interna para una escala unidimensional. Otros autores como Taber (2017), entran en más detalle sobre estos valores, indicando que diferentes autores interpretan los valores del Alfa de Cronbach de la siguiente manera: puntajes entre 0.61 - 0.65 se consideran moderados, puntajes entre 0.65 - 0.70 se consideran adecuados, puntajes entre 0.70-0.76 se consideran relativamente altos, puntajes entre 0.76-0.81 se consideran bastante altos, puntajes entre 0.81-

0.84 se consideran robustos, puntajes entre 0.84-90 se consideran confiables, puntajes entre 0.91-0.93 se consideran fuertes, y puntajes entre 0.93-0.94 se consideran excelentes.

15.7. Funcionamiento de las categorías de respuesta de los ítems

Finalmente, en lo que respecta a los análisis con el Modelo de Rasch, es importante tener claro que la extensión del modelo de Rasch de datos dicotómicos hacia un modelo de escalas de calificación (RSM) ha tenido una gran influencia en la investigación, por lo que el RSM se utiliza hoy en día de forma rutinaria para el análisis de datos obtenidos con escalas tipo Likert, lo que según Bond y Fox (2007) representa una propuesta más apropiada a nivel psicométrico que lo que permitía el estilo tradicional de análisis de este tipo de datos.

Asimismo, Prieto y Delgado (2007) han señalado que al ser el RSM un modelo tipo Rasch para ítems politómicos, cuenta con propiedades métricas óptimas, además de que desde un punto de vista práctico tiene la ventaja de que no se requieren muestras muy grandes para lograr una estimación de parámetros apropiada. Por su parte, Linacre (2002) ha establecido una serie de reglas para evaluar el funcionamiento de una escala tipo Likert: se requieren al menos 10 observaciones para cada categoría de respuesta, debe existir una distribución regular de las observaciones a través de las categorías, se tiene que dar un incremento monotonóico de los promedios de las categorías y de los pasos entre las categorías, así como un ajuste apropiado en las categorías. Este modelo especifica que un conjunto de ítems comparte la misma estructura de calificación de la escala y proporciona mediciones promedio y umbrales de categorías para el instrumento completo.

15.8. Análisis de validez

15.8.1. Análisis Factorial Factorial Exploratorio

El Análisis Factorial Exploratorio, es uno de los análisis de ítems más utilizados del mundo, se lleva a cabo principalmente en estudios relacionados con la fabricación y validación de tests, ya que es el análisis por excelencia para identificar el constructo subyacente que existe sobre el conjunto de variables y/o factores que explican las respuestas de los ítems de un test (Lloret et tal., 2014). Es decir, para este test, al ser basado en el modelo de CAPP, mismo que se divide en seis dimensiones, lo esperado sería que el Análisis Factorial Exploratorio brinde seis factores, y que, en estos, se ajusten de manera adecuada las escalas pertenecientes a cada dominio.

15.8.2. Análisis T Student

Además de los análisis anteriores, se realizó una prueba t para determinar el promedio de media de dos subgrupos de la muestra, se busca determinar posibles diferencias significativas en las puntuaciones que puedan ser determinadas por la variable género.

15.8.3. ANOVA

El Análisis de Varianza (ANOVA), según Ritchey (2008) se puede considerar una extensión del análisis T. Ya que mediante este se pone a prueba si las medias de dos o más grupos son iguales, esto al ponerlas bajo análisis en conjunto con otros factores que pueden afectar dicha media. Se parte del hecho de que las medias son iguales para todos los grupos, esta sería la hipótesis nula, mientras que la hipótesis alterna indicaría que hay al menos un grupo cuya media fue alterada por alguno de los factores del estudio.

16. Resumen metodológico

Tabla 2

Resumen metodológico

Fases metodológicas	Pasos a seguir o breve descripción
Especificaciones del estudio	En la identificación del propósito y finalidad del <i>test</i> , se delimitan alcances y usos de la prueba a desarrollar. En las especificaciones para la aplicación, se determina como una prueba de auto reporte y la necesidad de un nivel mínimo de lectura de sexto grado de primaria.

Descripción de la población y muestra	<p>La muestra para este proyecto serán hombres y mujeres mayores de 18 años con un nivel de lectura mínimo de sexto año de primaria, se incluirá población forense y general.</p> <p>El método de muestreo utilizado fue a conveniencia debido a la limitante representada por la pandemia mundial por Covid-19 (ver apartado Limitaciones).</p>
Confidencialidad y protección a los participantes	Se brindará consentimientos informados a las personas participantes.
Retribución a la población	Se compartirán los resultados de la investigación, describiendo las características de la población y los principales resultados del estudio.
Diseño y creación de <i>ítems</i>	<p>Diseño de los <i>ítems</i> en base a los modelos teóricos propuestos en el CAPP (Cooke, et al., 2012), a lo largo de la dimensionalidad del constructo psicopático.</p> <p>Los <i>ítems</i> diseñados serán de tipo Likert en escala de 1 a 4.</p>
Selección y revisión de <i>ítems</i>	Entrevista cognitiva para verificar la comprensión de los <i>ítems</i> .

Ensamble de la prueba	<p>Paso 1: Selección de <i>ítems</i> y generación de escalas en base a la depuración en la etapa anterior, estos se van a agrupar en base a la representatividad que posean según las variables que se pretenden evaluar. Se modificarán <i>ítems</i> similares o que sugieran respuestas de otras preguntas en la prueba.</p> <p>Paso 2: Revisión de la prueba con expertos en la materia.</p>
Desarrollo del protocolo de aplicación	Construcción de instrucciones y protocolo para la aplicación de la prueba.
Estudio piloto	<p>Se realizó un pilotaje de la prueba ensamblada con grupo representativo de 60 personas de la muestra general. Los casos que presentaron resultados adecuados, fueron considerados parte de la muestra final.</p> <p>Revisión inicial de ítems mediante método Rasch para depurar el instrumento.</p>
Aplicación de la prueba	<p>Se aplicó la prueba construida.</p> <p>Población general: 432 personas</p> <p>Población forense: 18 personas</p>
Evidencias de validez de criterio	<p>Aplicación adicional de pruebas para obtención de evidencias de validez de criterio.</p> <p>Población general: Se aplicó el MMPI-2 (N:59 - 60), el IPPD (N: 70-71), PAI (N: 58), NEO PI-R (N: 59 - 60) y el UPPS-P (N: 140-141).</p>

Análisis de datos

Aplicación de análisis estadísticos con el modelo de Rasch (Infit y Outfit). Realización de análisis de funcionamiento diferencial de los ítems (DIF), Análisis de componentes principales de los residuos (ACPR). Igualmente se llevarán a cabo análisis de alfa de Cronbach, análisis factorial exploratorio (AFE), análisis de T de Student, ANOVA, y coeficiente de Pearson.

Fuente: elaboración propia

Resultados

1. Análisis de Rasch

En el siguiente apartado, se elaborará una explicación por escala, la cual brindará con detalles los resultados de los análisis iniciales de cada escala con la totalidad de los ítems administrados. A continuación, se indican los ítems que no se ajustan al modelo de Rasch (Infit y Outfit), los ítems que presentaron una correlación ítem-total baja (RiX), razones por las que pudieron ser eliminados.

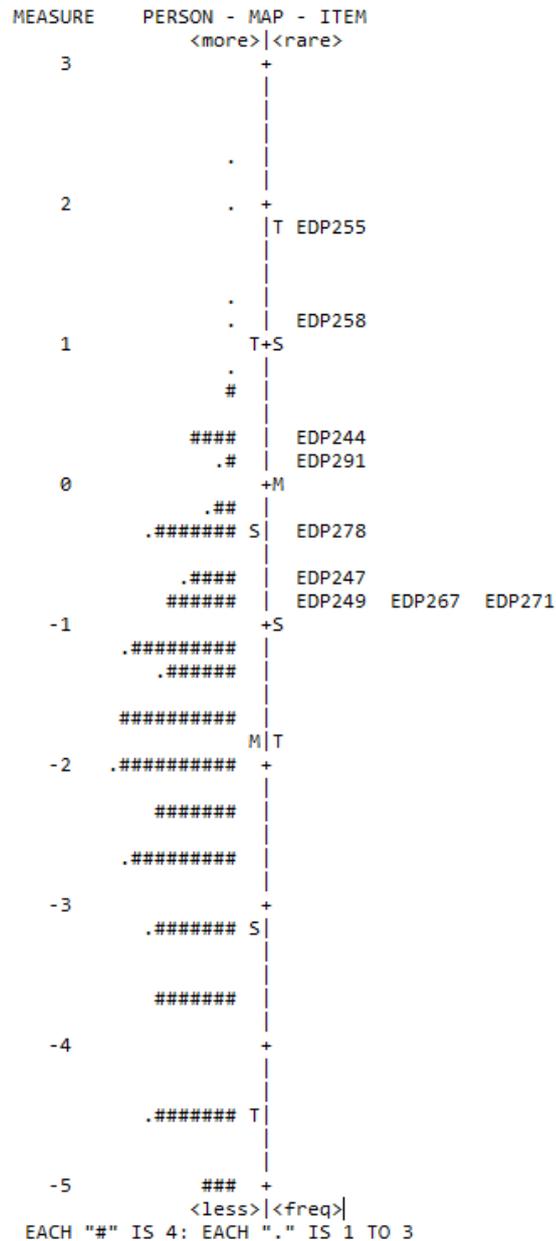
Se señala, además, el Análisis de Componentes Principales de los residuos, indicador de una segunda dimensión, lo que puede indicar la necesidad de eliminar algún ítem. Asimismo, se reportan los ítems que presentaron funcionamiento diferencial (DIF mayor a .05). Asimismo, se reportan los indicadores de confiabilidad y separación de los ítems y las personas en este análisis final de la escala.

1.1. Dominio de Sí Mismo

1.1.1. Escala Centrado en mí mismo

Figura 2

Gráfico Escala Centrado en Sí Mismo – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Centrado en Sí Mismo originalmente constaba de 10 ítems, durante el proceso de ajuste se eliminó un ítem. El ítem 275 desajustó severamente el infit y el outfit (infit=2.06, outfit = 2.36), por lo que se tomó la decisión de eliminarlo de la escala.

Finalmente, siguiendo con los criterios de Linacre (2022), ningún ítem presentó funcionamiento diferencial, correlación punto-media baja o afectaciones a la dimensionalidad. De esta forma, la escala de Centrado en Sí Mismo quedó en definitiva con nueve ítems en total.

Resultados finales de la escala

Tabla 3

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Centrado en Sí Mismo.

	Ítems	Personas
Fiabilidad	.99	.79
Separación	10.91	1.94

Fuente: elaboración propia

Tal y como se observa en la Tabla 3, los estadísticos de fiabilidad y separación indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos congruentes a los síntomas de la escala Centrado en Sí Mismo y que podríamos esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. De igual manera, la escala presenta un alfa de Cronbach de .82.

Tabla 4

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Centrado en Sí Mismo.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1674	43	-3.01	.97	.97	No
2	1457	37	-1.28	.95	.78	-2.04
3	637	16	-.20	.99	1.00	.06
4	120	3	.48	1.35	1.42	1.98

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Centrado en Sí Mismo cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas la escala.

Tabla 5

Estadísticos de los ítems para la escala Centrado en Sí Mismo.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM267	-.82	.08	.87	.88	-2.00	-1.90	.69	.11
ITEM271	-.86	.08	.93	.91	-1.07	-1.37	.70	.07
ITEM291	.21	.08	.96	.92	-.55	-1.00	.66	-.16
ITEM255	1.87	.11	1.02	.79	.28	-1.33	.51	.43
ITEM258	1.19	.10	1.03	.77	.41	-2.03	.60	.41
ITEM278	-.39	.08	1.25	1.25	3.44	3.31	.56	-.12
ITEM244	.39	.09	1.25	1.14	3.28	1.53	.61	.00
ITEM247	-.67	.08	.80	.81	-3.17	-2.94	.69	-.07
ITEM249	-.91	.08	1.13	1.18	1.96	2.57	.55	-.31

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Centrado en Sí Mismo indican que, en el caso de los ítems, el promedio del Infit fue de 1.03 (DT= .15). Para el ajuste global externo, el promedio del Outfit fue de .96 (DT=.17). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT= .60) y el valor del Outfit de .96 (DT= .61). De esta forma, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 9.49% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias

significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.51 a 0.70) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 47.7%. Asimismo, el primer componente tenía un valor propio de 1.96 lo que corresponde a un 11.4% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

Los nueve ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala Centrado en Sí Mismo que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.1.2. Escala Autoagrandecimiento

Figura 3

Gráfico Escala Autoagrandecimiento – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Autoagrandecimiento constaba inicialmente de 9 ítems, durante el proceso de análisis y ajuste se eliminó el ítem 1. Esto debido a que podría estar presentando funcionamiento diferencial (DIF contrast = 0.80, $p < .001$) siendo de menor severidad para los hombres de la muestra. Ningún ítem presentó alteraciones de correlación punto medida o afectaciones a la unidimensionalidad (porcentaje de varianza explicada fue de 44.2% con un primer contraste pasó de 1.79). La escala de Autoagrandecimiento quedó en definitiva con ocho ítems en total.

Resultados finales de la escala

Tabla 6

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Autoagrandecimiento.

	Ítems	Personas
Fiabilidad	.99	.69
Separación	10.63	1.49

Fuente: elaboración propia

En la Tabla 6 se muestran los estadísticos de fiabilidad y separación, que indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos de autoagrandecimiento y que podríamos esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. En cuanto a la confiabilidad según el alfa de Cronbach, esta escala presentó .73.

Tabla 7

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Autoagrandecimiento.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1141	33	-2.37	.96	.97	No
2	1505	44	-1.03	1.00	.87	-1.98
3	625	18	-.04	1.01	1.00	.31
4	185	5	.76	1.19	1.20	1.67

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Autoagrandecimiento cumple con los cuatro criterios propuestos por Linacre (2002), y se garantizan la calidad de las categorías de respuesta utilizadas.

Tabla 8

Estadísticos de los ítems para la escala Autoagrandecimiento.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM297	1.40	.09	.88	.78	-1.69	-2.50	.64	.57
ITEM297	.87	.08	.97	.91	-.36	-1.18	.61	.07
ITEM251	.06	.08	.97	.96	-.43	-.62	.69	.31
ITEM4	.44	.08	1.00	1.01	.05	.18	.53	.25
ITEM7	-.26	.07	.91	.92	-1.33	-1.27	.53	.00
ITEM337	-.53	.07	1.15	1.13	2.13	1.92	.62	-.46
ITEM340	-1.50	.07	1.08	1.07	1.24	1.10	.50	-.34
ITEM10	-.48	.07	1.13	1.11	1.95	1.68	.52	-.03

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Autoagrandecimiento, para los ítems, el promedio del Infit fue de 1.01 (DT=.09), y para el ajuste global externo, el promedio del Outfit fue de .99 (DT=.11). Mientras el valor del ajuste interno obtenido fue de 1.01 (DT=.60) y el valor del Outfit de .99 (DT= .58) en los indicadores para personas. De esta manera, se evidencia que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), y se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

A su vez, solamente 7.4% de las personas presentaron valores mayores a 2 en Infit u Outfit, lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias

significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.50 a 0.69) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

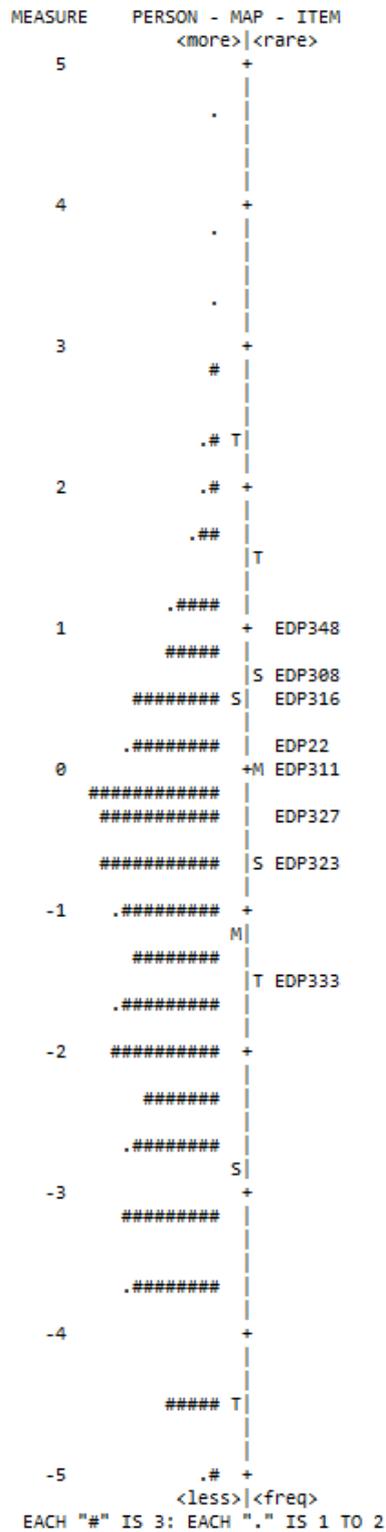
Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 44.2%, con un primer componente de valor propio de 1.79 lo que corresponde a un 12.5% de la varianza explicada. Por esto, la escala cumple con los criterios de unidimensionalidad establecidos.

En cuanto al DIF, el ítem 297, a pesar de tener posibilidad de presentar un leve funcionamiento diferencial (DIF contrast = 0.57, $p < .001$), se decide mantener en la escala por un criterio de contenido; ninguno otro ítem tiene un DIF por encima de los criterios establecidos para detectar funcionamiento diferencial. Además, la validez de constructo fue respaldada por todos los ítems en la escala Autoagrandecimiento que tenían un buen ajuste bajo el modelo de Rasch y correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.1.3. Escala Sentido de Ser Único

Figura 4

Gráfico Escala Sentido de Ser Único – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Sentido de Ser Único estaba originalmente compuesta de 10 ítems. En un primer análisis, el ítem 17 presentó una elevación en los valores de infit (1.53) y outfit (1.52), por lo que se eliminó de la escala. Tras un segundo análisis, el ítem 14 presentó indicios de funcionamiento diferencial (DIF=1.55), por lo que se decide eliminar también.

Los ítems presentaron un ajuste adecuado en cuanto a la correlación punto media y no hubo indicios de afectación a la unidimensionalidad. De esta manera, la escala de Sentido de Ser Único quedó en definitiva con ocho ítems en total.

Resultados finales de la escala

Tabla 9

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Sentido de Ser Único.

	Ítems	Personas
Fiabilidad	.99	.86
Separación	9.02	2.46

Fuente: elaboración propia

En la Tabla 9 se muestran los estadísticos de fiabilidad y separación, que indican que la escala es eficiente para diferenciar a personas con diferentes niveles de rasgos alineados con la escala de sentido de ser único, se pueden esperar consistencia de estas inferencias; además, se comprueba que existe una jerarquía en los ítems de la escala. .86, también, fue la medida de alfa de Cronbach en esta escala, reforzando esta confiabilidad.

Tabla 10

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Sentido de Ser Único.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	979	28	-2.89	1.02	1.04	No
2	1315	38	-1.20	.93	.88	-2.38
3	970	28	.18	1.02	1.05	-.21
4	192	6	1.87	1.03	1.03	2.59

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Sentido de Ser Único cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta.

Tabla 11

Estadísticos de los ítems para la escala Sentido de Ser Único.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM323	-.60	.08	.93	1.01	-1.07	.17	.66	-.49
ITEM327	-.41	.08	.96	.95	-.66	-.68	.73	-.32
ITEM348	1.08	.09	.88	.81	-1.86	-2.23	.72	1.02
ITEM308	.72	.08	.81	.73	-3.04	-3.68	.78	.57
ITEM311	.08	.08	.76	.72	-3.86	-4.35	.80	.03
ITEM316	.44	.08	.98	.96	-.33	-.56	.71	.14
ITEM333	-1.43	.08	1.44	1.41	5.82	5.28	.68	-1.19
ITEM22	.12	.08	1.22	1.35	3.18	4.49	.59	.23

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Sentido de Ser Único indican un promedio del Infit de 1.00 (DT=.21), para los ítems. Respecto al ajuste global externo, el promedio del Outfit fue de .99 (DT=.24). A su vez, los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT=.59) y el valor del Outfit de .99 (DT= .59). Por consiguiente, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se concluye que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente el síntoma.

Por otro lado, 7.17% de las personas presentaron valores mayores a 2 en Infit u Outfit, un porcentaje adecuado que indica que el ajuste de las personas al modelo es aceptable. Los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los

datos empíricos obtenidos. A su vez, el análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.59 a 0.80) por lo que se evidencia cumplen con el criterio de validez de contenido requerido.

En el abordaje del ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR), resultando en un porcentaje de la varianza explicada por la medida Rasch del 54.9%. Más específicamente, el primer componente tenía un valor propio de 1.55 lo que corresponde a un 8.8% de la varianza explicada, cumpliendo con los criterios de unidimensionalidad establecidos.

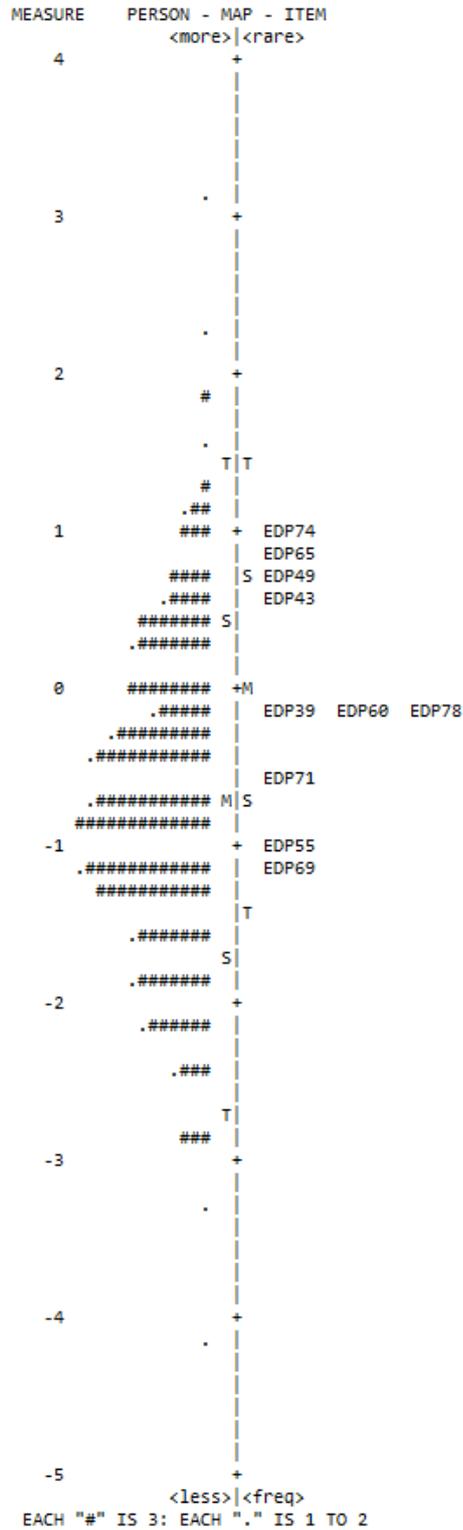
De los ocho ítems de la escala, los reactivos 348, 308 y 333 tienen un DIF por encima de los criterios establecidos para detectar dicho criterio, presentando la posibilidad de funcionamiento diferencial. Sin embargo, a pesar de estos valores, por una cuestión de contenido y porque la escala podría resultar siendo muy pequeña y afectar criterios de confiabilidad, se decide mantener estos ítems teniendo en cuenta este posible sesgo, ya que consideramos que es necesario analizarlos con una muestra más amplia y heterogénea antes de tomar una decisión definitiva sobre si deben ser eliminados.

A su vez, la validez de constructo fue respaldada por todos los ítems en la escala Sentido de Ser Único, con un buen ajuste bajo el modelo de Rasch, libres de DIF y con correlaciones ítem-total positivas y cercanas a sus valores esperados.

1.1.4. Escala Sentido de Derecho

Figura 5

Gráfico Escala Sentido de Derecho – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Sentido de Derecho constaba inicialmente de 11 ítems, durante el proceso de análisis y ajuste se eliminó el ítem 64. En un primer análisis, los ítems 64, 71, 55, y 78, desajustaron severamente el infit y/o outfit, con base en este análisis, se elimina el ítem 64, al ser el que presentó mayor desajuste (infit=4.94, outfit=8.74); una vez eliminado, se realiza un segundo análisis donde los demás ítems mejoraron sus medidas significativamente y no presentaron más alteraciones por las que ameritaran ser eliminados. Ningún ítem presenta funcionamiento diferencial, o afectaba la unidimensionalidad de la escala.

Por otro lado, los ítems presentaron una correlación punto medida adecuada. De esta manera, la escala de Sentido de Derecho quedó en definitiva con diez ítems en total.

Resultados finales de la escala

Tabla 12

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Sentido de Derecho.

	Ítems	Personas
Fiabilidad	.99	.79
Separación	10.11	1.94

Fuente: elaboración propia

En la Tabla 12 se muestran los estadísticos de fiabilidad y separación. Esta indica que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos alineados a los síntomas de sentido de derecho, por lo que se podría esperar consistencia de estas inferencias; además, se comprueba que existe una jerarquía en los ítems de la escala. Apoyando estos resultados, el alfa de Cronbach de la escala es de .78.

Tabla 13

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Sentido de Derecho.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1129	26	.53	.94	.96	-6.93
2	1780	41	1.61	.99	.91	.63

3	1011	23	2.42	.97	.96	2.54
4	399	9	3.15	1.22	1.22	3.76

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Sentido de Derecho cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 14

Estadísticos de los ítems para la escala Sentido de Derecho.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM69	-1.14	.07	1.02	1.03	.40	.49	.58	-.10
ITEM71	-.58	.07	.91	.93	-1.38	-1.16	.58	-.27
ITEM74	1.01	.08	.99	.97	-.14	-.34	.62	-.26
ITEM55	-.93	.07	.87	.86	-2.19	-2.26	.54	.00
ITEM60	-.15	.07	1.08	1.11	1.24	1.67	.43	.28
ITEM78	-.14	.07	.93	.93	-1.11	-1.03	.56	.15
ITEM39	-.17	.07	1.18	1.16	2.64	2.36	.60	-.10
ITEM43	.63	.07	1.05	.99	.78	-.07	.61	.24
ITEM49	.66	.07	1.07	1.03	1.06	.46	.62	.05
ITEM65	.80	.08	1.06	1.00	.83	.01	.58	.00

Fuente: Elaboración propia

Según los estadísticos de ajuste respectivos para la escala Sentido de Derecho, en el caso de los ítems, el promedio del Infit fue de 1.02 (DT=.09), mientras que, para el ajuste global externo, el promedio del Outfit fue de 1.00 (DT=.08). Por otro lado, en los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT=.58) y el valor del Outfit de 1.00 (DT= .59), donde los ítems presentan índices de ajuste adecuados (Infit u Outfit >

1.5), evidenciando que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

El porcentaje de las personas que presentaron valores mayores a 2 en Infit u Outfit fue de solamente 6.94%, determinando que el ajuste de las personas al modelo es aceptable. Por tanto, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.43 a 0.62) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

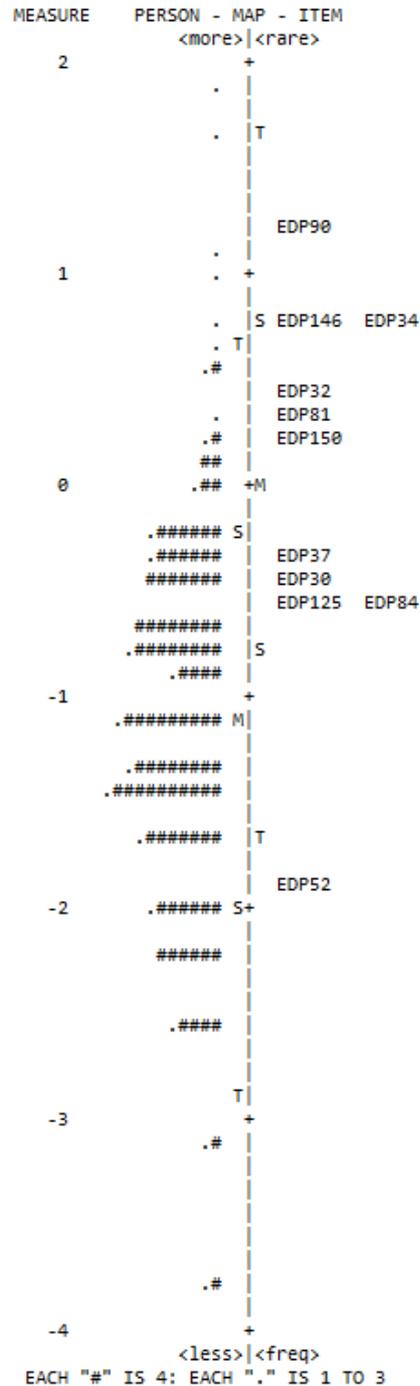
Para abordar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR), este análisis mostro que el porcentaje de la varianza explicada por la medida Rasch fue del 42.5%. Asimismo, el primer componente tenía un valor propio de 1.88 lo que corresponde a un 10.8% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

Los diez ítems restantes muestran un buen funcionamiento diferencial, ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo. De igual forma, la validez de constructo fue respaldada por todos los ítems en la escala Sentido de Derecho que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.1.5. Escala Sentido de Invulnerabilidad

Figura 6

Gráfico Escala Sentido de Invulnerabilidad – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Sentido de Invulnerabilidad constaba inicialmente de 12 ítems, durante el proceso de análisis y ajuste se eliminó el ítem 98. Este reactivo presentó elevación en el

valor de outfit (2.02), por lo que se eliminó de la escala. Ningún ítem presenta funcionamiento diferencial, o afectaba la unidimensionalidad de la escala. De igual forma, no presentaron alteraciones en la correlación punto media. La escala de Sentido de Invulnerabilidad quedó en definitiva con once ítems en total.

Resultados finales de la escala

Tabla 15

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Sentido de Invulnerabilidad.

	Ítems	Personas
Fiabilidad	.99	.73
Separación	11.53	1.65

Fuente: elaboración propia

En la Tabla 15 se muestran los estadísticos de fiabilidad y separación, que indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos de sentido de invulnerabilidad. Se podría esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. Además, el alfa de Cronbach de esta escala es .72, lo que apoya los resultados.

Tabla 16

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Sentido de Invulnerabilidad.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2034	43	-1.86	.97	.99	No
2	1525	32	-.95	.97	.94	-1.11
3	850	18	-.13	.92	.89	-.02
4	343	7	.49	1.20	1.24	1.12

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Sentido de Invulnerabilidad cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 17

Estadísticos de los ítems para la escala Sentido de invulnerabilidad.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM30	-.43	.06	.58	.61	-7.97	-6.80	.59	-.17
ITEM32	.45	.07	.74	.71	-4.03	-3.73	.60	.44
ITEM34	.74	.08	1.31	1.28	3.63	2.73	.38	.00
ITEM52	-1.88	.06	1.11	1.16	1.74	2.29	.49	-.02
ITEM81	.28	.07	1.22	1.19	3.03	2.23	.51	.31
ITEM84	-.51	.06	1.23	1.32	3.46	4.40	.42	-.47
ITEM90	1.23	.09	.99	.86	-.10	-1.17	.46	.00
ITEM37	-.28	.06	1.15	1.15	2.31	2.15	.43	-.09
ITEM146	.75	.08	.79	.77	-2.86	-2.63	.55	.06
ITEM150	.19	.07	.90	.84	-1.53	-2.13	.59	.12
ITEM125	-.54	.06	1.05	1.03	.84	.53	.57	.00

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Sentido de Invulnerabilidad indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT=.22). Para el ajuste global externo, el promedio del Outfit fue de .99 (DT=.23). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT=.56) y el valor del Outfit de .99 (DT= .56). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 6.71% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las

personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.38 a 0.60) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 43.4%. Asimismo, el primer componente tenía un valor propio de 1.84 lo que corresponde a un 9.5% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

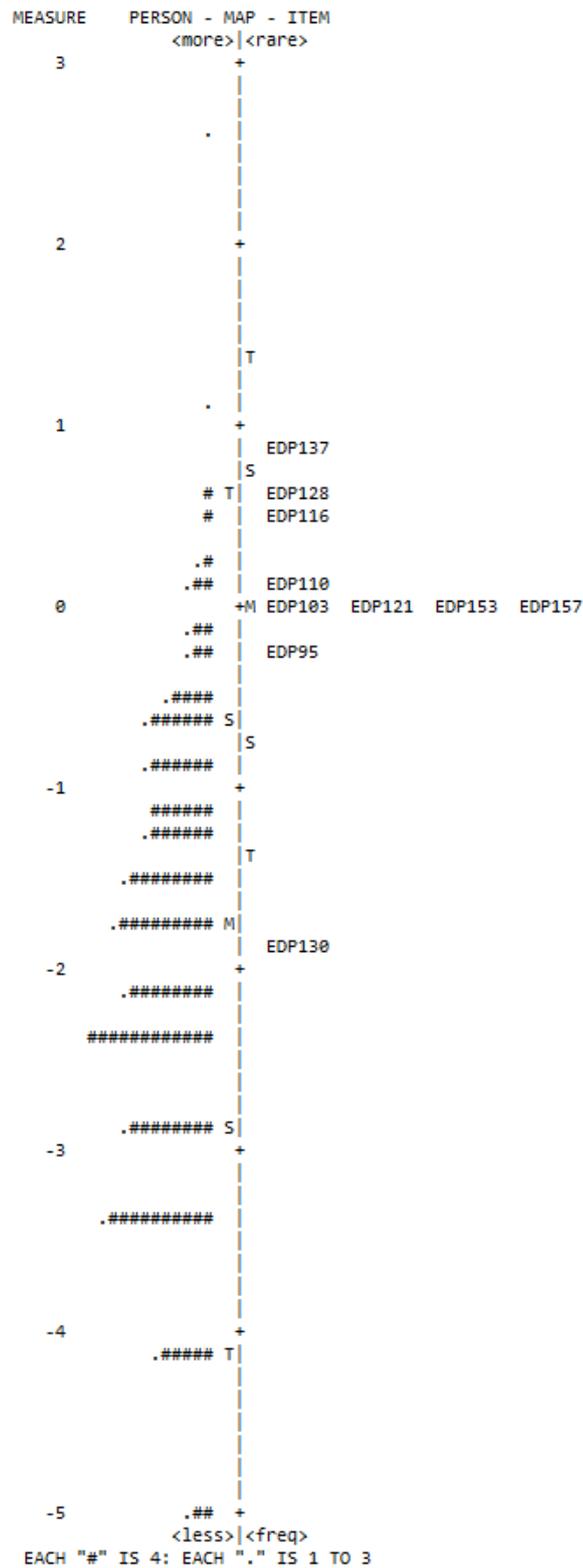
Los once ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala Sentido de Invulnerabilidad que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.1.6. Escala Autojustificación

Figura 7

Gráfico Escala Autojustificación – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Autojustificación constaba inicialmente de 11 ítems, durante el proceso de análisis y ajuste se eliminó un ítem. Durante este proceso, se identificó que el ítem 161 estaba muy cerca del punto de corte para el valor de el infit (1.32) y outfit (1.46), adicionalmente, poseía una baja correlación punto corte $RiX = .37$. Con base en estos criterios y un análisis de contenido, se decide eliminar el ítem, lo cual mejora un poco la fiabilidad (.74 a .75) y según el Análisis de componentes principales de los residuos (ACPR), el porcentaje de la varianza explicada por la medida Rasch incrementó de 39.7% a 42.6%.

Ningún otro ítem presentó funcionamiento diferencial, alteraciones en la correlación punto media, o afectaba la unidimensionalidad de la escala. De esta manera, la escala de Autojustificación quedó en definitiva con diez ítems en total.

Resultados finales de la escala

Tabla 18

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Autojustificación.

	Ítems	Personas
Fiabilidad	.99	.75
Separación	8.42	1.71

Fuente: elaboración propia

En la Tabla 18 se muestran los estadísticos de fiabilidad y separación. Estos indican que la escala es eficiente para diferenciar a personas con diferentes niveles de rasgos de autojustificación, por lo que se podría esperar consistencia de estas inferencias y una jerarquía en los ítems de la escala. Lo cual es apoyado por el alfa de Cronbach de .79.

Tabla 19

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Autojustificación.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2139	50	-2.56	1.01	1.00	No
2	1548	36	-1.36	.98	.91	-1.69

3	495	11	-.32	.93	.89	.24
4	138	3	.47	1.30	1.46	1.45

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Autojustificación cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 20

Estadísticos de los ítems para la escala Autojustificación.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM128	.65	.09	.82	.65	-2.39	-3.98	.65	.26
ITEM130	-1.88	.07	1.14	1.14	2.14	2.08	.56	-.41
ITEM153	.05	.08	.92	.90	-1.12	-1.22	.61	-.15
ITEM110	.09	.08	1.13	1.28	1.77	3.17	.44	.34
ITEM116	.44	.09	.92	.87	-1.07	-1.45	.61	.13
ITEM137	.82	.09	1.02	1.13	.29	1.20	.43	.19
ITEM95	-.20	.08	1.04	1.03	.57	.41	.59	-.26
ITEM103	-.04	.08	.96	.93	-.56	-.84	.55	.13
ITEM121	.05	.08	1.16	1.06	2.11	.69	.57	.07
ITEM157	.00	.08	1.00	.91	.05	-1.15	.63	.00

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Autojustificación indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT=.11). Para el ajuste global externo, el promedio del Outfit fue de .99 (DT=.17). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT=.58) y el valor del Outfit de .99 (DT= .59). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit >

1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 6.01% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.43 a 0.65) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 42.6%. Asimismo, el primer componente tenía un valor propio de 1.74 lo que corresponde a un 10% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

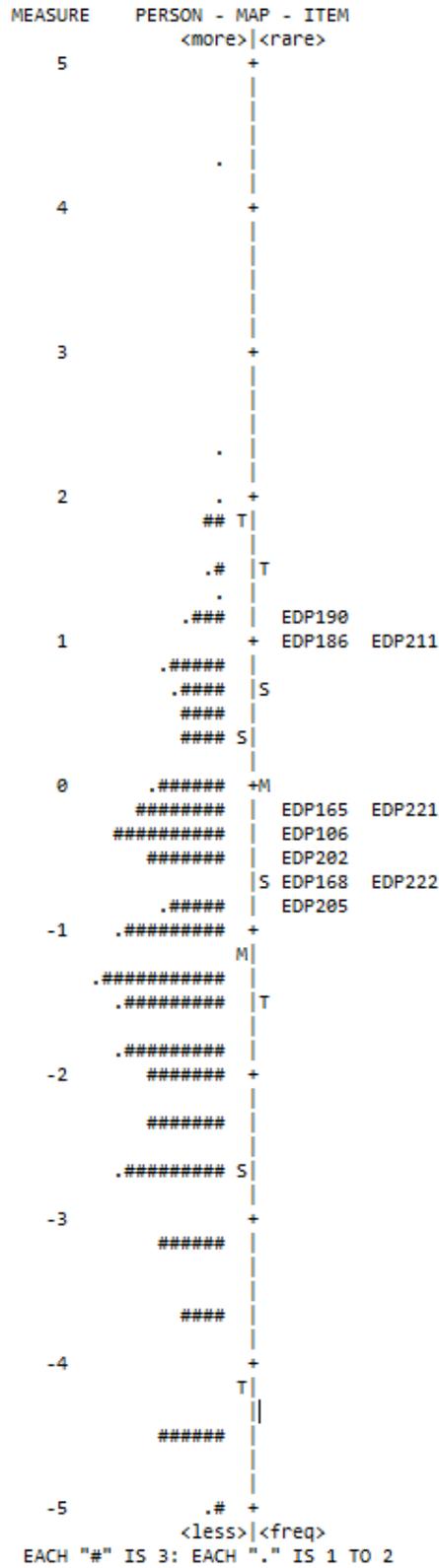
Los diez ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala Autojustificación que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.1.7. Escala Autoconcepto Inestable

Figura 8

Gráfico Escala Autoconcepto Inestable – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Autoconcepto Inestable constaba inicialmente de 11 ítems, durante el proceso de análisis y ajuste se eliminó un ítem. No hubieron alteraciones en cuanto al infit u outfit, pero el ítem 209, siguiendo con los criterios de Linacre (2022), podría estar presentando funcionamiento diferencial, por lo que se tomó la decisión de eliminar el (DIF contrast =-.56, $p < .001$) siendo de menor severidad para las mujeres de la muestra. Asimismo, el ítem 209 afectaba la unidimensionalidad, al ser eliminado el porcentaje de varianza explicada pasó de 49.3 % a 50.7%, además, el primer contraste pasó de 2.04 a 1.96.

Los ítems no presentaron alteraciones en la medida punto media. Con base en esto, la escala de Autoconcepto Inestable quedó en definitiva con diez ítems en total.

Resultados finales de la escala

Tabla 21

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Autoconcepto Inestable.

	Ítems	Personas
Fiabilidad	.99	.86
Separación	8.97	2.45

Fuente: elaboración propia

En la Tabla 21 se muestran los estadísticos de fiabilidad y separación, indicando que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos de un autoconcepto inestable, con una consistencia de estas inferencias. Además, el con un valor de alfa de Cronbach de .87, se reafirma esta conclusión, y que se comprueba que existe una jerarquía en los ítems de la escala.

Tabla 22

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Autoconcepto Inestable.

Categoría	Muestra	Frecuencia	Promedio	InFit	OutFit	Paso
		(%)				
1	1364	32	-2.57	1.04	1.03	No
2	1788	41	-1.09	.91	.93	-2.08
3	895	21	.23	.87	.84	.26
4	273	6	1.05	1.17	1.25	1.83

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Autoconcepto Inestable cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 23

Estadísticos de los ítems para la escala Autoconcepto Inestable.

Ítem	Nivel de	Error	MNSQ		ZTSD		Correlacione	DIF
			severidad	estándar	Infit	Outfit		
	(en						medida	
	logits)							
ITEM165	-.16	.08	1.24	1.27	3.34	3.51	.58	.15
ITEM106	-.31	.08	.93	.96	-1.08	-.55	.61	.00
ITEM221	-.19	.08	1.04	1.01	.65	.15	.70	.08
ITEM168	-.66	.07	.78	.79	-3.63	-3.42	.75	-.23
ITEM202	-.47	.08	.85	.86	-2.36	-2.19	.75	-.16
ITEM205	-.76	.07	.95	.98	-.80	-.30	.72	.10
ITEM211	1.02	.09	.89	.80	-1.62	-2.32	.67	.28
ITEM222	-.60	.07	.93	.91	-1.11	-1.41	.70	-.27
ITEM186	1.02	.09	1.14	1.05	1.92	.56	.60	.07
ITEM190	1.12	.09	1.31	1.29	3.86	2.74	.50	.09

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Autoconcepto inestable indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT=.16). Para el ajuste global externo, el promedio del Outfit fue de .99 (DT=.17). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de .99 (DT=.61) y el valor del Outfit de .99 (DT= .61). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 8.8% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.50 a 0.75) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 50.7%. Asimismo, el primer componente tenía un valor propio de 1.96 lo que corresponde a un 9.7% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

Los diez ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

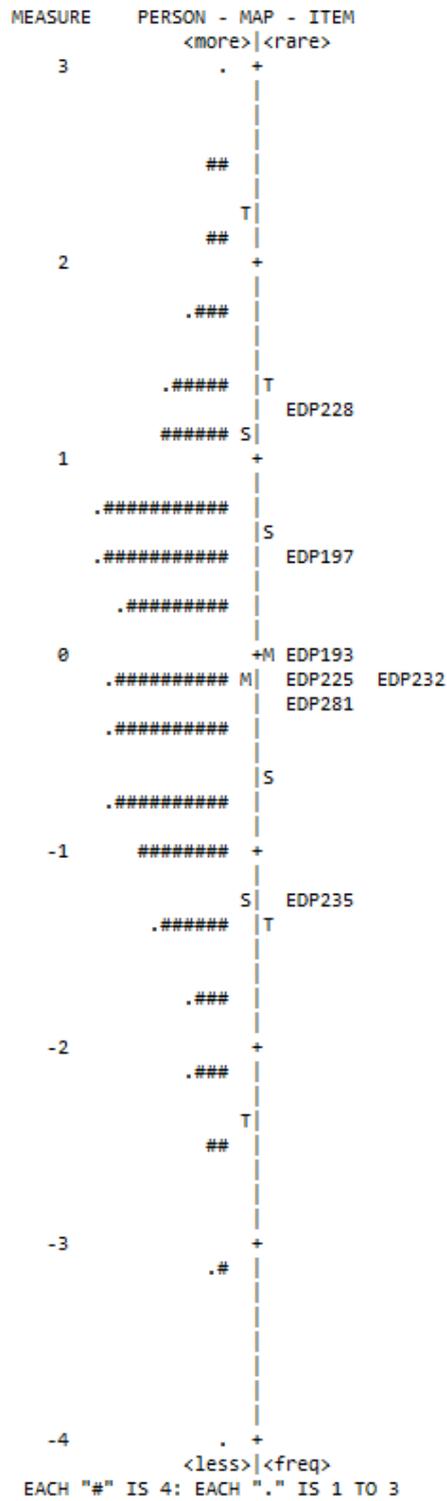
La validez de constructo fue respaldada por todos los ítems en la escala Autoconcepto Inestable que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.2. Dominio Emocional

1.2.1. Falta de Ansiedad

Figura 9

Gráfico Escala Falta de Ansiedad – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Falta de Ansiedad estaba conformada de 10 ítems, durante el proceso de ajuste se eliminaron tres ítems. El ítem 214 desajustó en los valores de infit (1.58) y outfit (1.51); además, tanto el 214 como los ítems 179 y 285 presentaron indicios de funcionamiento diferencial (DIF contrast = 0.80, $p < .001$), siendo de mayor severidad para las mujeres de la muestra.

Ningún ítem presentó alteraciones de correlación punto medida, ni tampoco afectaciones a la unidimensionalidad, según el cual, el porcentaje de varianza explicada fue de 55.2% con un primer contraste pasó de 1.76. La escala de Falta de Ansiedad quedó en definitiva con siete ítems en total.

Resultados finales de la escala

Tabla 24

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Falta de Ansiedad.

	Ítems	Personas
Fiabilidad	.99	.76
Separación	9.41	1.77

Fuente: elaboración propia

En la Tabla 24 se muestran los estadísticos de fiabilidad y separación, adicionalmente, con un alfa de Cronbach de .72, se indica que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos de ausencia de ansiedad, y que se podría esperar buena consistencia de estas inferencias, así una jerarquía en los ítems de la escala.

Tabla 25

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Falta de Ansiedad.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	447	15	-1.47	1.15	1.13	(-3.17)
2	1170	39	-.56	.92	.95	-.97
3	1001	33	.49	.90	.90	1.04
4	406	13	1.44	.97	1.83	(3.05)

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Falta de Ansiedad cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 26

Estadísticos de los ítems para la escala Falta de Ansiedad.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM193	-.02	.07	1.01	1.02	.16	.27	.67	.45
ITEM197	.50	.07	.85	.85	-2.40	-2.43	.61	-.13
ITEM225	-.13	.07	1.04	1.03	.59	.54	.65	.44
ITEM228	1.21	.08	1.20	1.17	2.81	2.33	.40	-.29
ITEM232	-.17	.07	.88	.88	-1.95	-1.91	.69	.14
ITEM235	-1.20	.07	1.07	1.08	1.17	1.25	.56	-.39
ITEM281	-.20	.07	.92	.92	-1.23	-1.19	.64	-.28

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Falta de Ansiedad indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT=.11). Para el ajuste global externo, el promedio del Outfit fue de .99 (DT=.11). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.00 (DT=.67) y el valor del Outfit de .99 (DT= .65). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 9.02% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.40 a 0.67) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

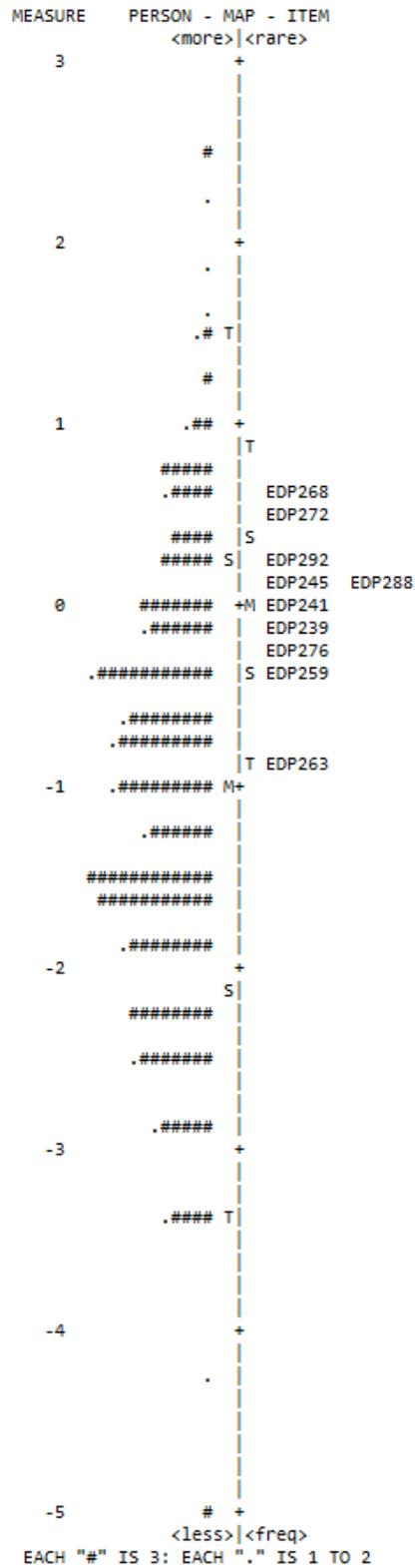
Mediante un Análisis de componentes principales de los residuos (ACPR), se investigó el ajuste de la prueba al supuesto de unidimensionalidad. El porcentaje de la varianza explicada por la medida Rasch fue del 55.2%. Asimismo, el primer componente tenía un valor propio de 1.76, un 13.9% de la varianza explicada, por lo que cumple con los criterios de unidimensionalidad establecidos.

Los siete ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo. Asimismo, la validez de constructo fue respaldada por todos los ítems en la escala Falta de Ansiedad que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.2.2. Escala No Experimental Placer

Figura 10

Gráfico Escala No Experimental Placer – Distribución casos e ítems



Resultados iniciales de la escala

La escala de No Experimenta Placer estaba conformada de 11 ítems, durante el proceso de análisis, se determinó la existencia de una segunda dimensión, por lo que un ítem estaba afectando la unidimensionalidad de la escala. A pesar de no presentar alteraciones en los valores de infit, outfit, correlación punto media, se decide eliminar el ítem 279; debido a indicaciones de funcionamiento diferencial ($DIF=.57$).

El Análisis de componentes principales de los residuos indicó, en un primer análisis, un porcentaje de varianza explicada de 41.5% y un primer contraste de 2.07, una vez eliminado el ítem 279, estos valores pasaron a 41.3% y 1.69 respectivamente. Ningún otro reactivo presentó alteraciones en sus valores de ajuste; de esta manera, la escala de No Experimenta Placer quedó en definitiva con diez ítems en total.

Resultados finales de la escala

Tabla 27

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala No Experimenta Placer.

	Ítems	Personas
Fiabilidad	.97	.83
Separación	5.59	2.17

Fuente: elaboración propia

Según los datos de la Tabla 27, los estadísticos de fiabilidad y separación, y con un alfa de Cronbach de .83, la escala es sensible para diferenciar a personas con diferentes niveles de rasgos asociados con la falta de experimentar placer. Además, se podría esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala.

Tabla 28

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala No Experimental Placer.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1249	29	-2.03	.99	1.00	No
2	1814	42	-.97	.96	.91	-1.89
3	1024	24	-.03	.92	.90	.03
4	233	5	.65	1.22	1.27	1.86

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala No Experimental Placer cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 29

Estadísticos de los ítems para la escala No Experimental Placer.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM288	.12	.07	1.05	1.08	.72	1.15	.56	-.30
ITEM239	-.10	.07	1.10	1.08	1.52	1.22	.56	-.21
ITEM268	.67	.08	1.03	.97	.43	-.34	.62	-.15
ITEM272	.50	.08	1.21	1.08	2.95	1.20	.72	.07
ITEM276	-.23	.07	.73	.76	-4.53	-4.02	.64	-.17
ITEM292	.26	.07	1.22	1.17	3.12	2.46	.61	.00
ITEM259	-.42	.07	.96	.95	-.69	-.74	.66	.06
ITEM263	-.86	.07	1.05	1.07	.77	1.17	.48	.00
ITEM241	-.02	.07	.81	.80	-3.04	-3.19	.71	.35
ITEM245	.08	.07	.97	.97	-.47	-.51	.64	.35

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala No Experimenta Placer indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT=.15). Para el ajuste global externo, el promedio del Outfit fue de .99 (DT=.12). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.00 (DT=.59) y el valor del Outfit de .99 (DT= .59). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 5.55% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.48 a 0.72) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR), indicando un porcentaje de la varianza explicada por la medida Rasch fue del 58.7%. El primer componente tenía un valor propio de 1.69 lo que corresponde a un 9.9% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

Los diez ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala No Experimenta Placer que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.2.3. Escala Falta de Profundidad Emocional

Figura 11

Gráfico Escala Falta de Profundidad Emocional – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Falta de Profundidad Emocional estaba conformada de 11 ítems. Durante el proceso de análisis, se identificaron problemas con los ítems 250, 301 y 8, por lo cual se

tomó la decisión de eliminarlos. En un primer análisis, el ítem 8 presentó indicios de funcionamiento diferencial (DIF=.80) siendo de menor severidad para los hombres de la muestra. Similarmente, el ítem 250 presentó una correlación punto media baja de $R_{iX}=.29$. Adicionalmente, en este primer análisis, se identificó una segunda dimensión con el análisis (porcentaje de varianza explicada 40.2%, primer contraste 2.25).

Una vez eliminados los ítems 8 y 250, tras un segundo análisis, siguió presentando una segunda dimensión (porcentaje de varianza explicada 44.5%, primer contraste 2.20), por lo cual se busca eliminar el ítem que peor ajuste en estos valores. En este caso, también se utiliza un criterio con base en el contenido y tipo de ítem (si es inverso), por lo cual se decide eliminar el 301. Un análisis final determinó una mejora en la unidimensionalidad (porcentaje de varianza explicada 47.5%, primer contraste 1.90). Por estas razones, la escala Falta de Profundidad Emocional se redujo a ocho ítems.

Resultados finales de la escala

Tabla 30

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Falta de Profundidad Emocional.

	Ítems	Personas
Fiabilidad	.99	.77
Separación	11.42	1.82

Fuente: elaboración propia

Con un alfa de Cronbach de .75 y los datos mostrados en la Tabla 30, estadísticos de fiabilidad y separación, se demuestra que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos impulsivos y puede esperar consistencia de estas inferencias. Asimismo, se comprueba que existe una jerarquía en los ítems de la escala.

Tabla 31

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Falta de Profundidad Emocional.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1001	29	-1.85	1.00	1.03	No
2	1184	34	-.80	.92	1.04	-1.45
3	922	27	.24	.83	.82	-.05
4	349	10	.80	1.18	1.20	1.50

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Falta de Profundidad Emocional cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 32

Estadísticos de los ítems para la escala Falta de Profundidad Emocional.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlaciones punto-medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM265	-1.05	.07	1.38	1.36	5.44	5.06	.56	.50
ITEM295	.56	.07	.70	.67	-5.12	-5.05	.74	-.08
ITEM298	.50	.07	.91	.93	-1.32	-.95	.61	-.10
ITEM252	-.54	.07	1.43	1.47	6.08	6.41	.39	-.19
ITEM5	1.57	.08	1.10	1.31	1.28	2.76	.36	.44
ITEM305	.08	.07	.79	.79	-3.55	-3.26	.72	.08
ITEM338	-.47	.07	.86	.86	-2.29	-2.33	.69	-.26
ITEM341	-.65	.07	.74	.75	-4.57	-4.31	.65	-.20

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Falta de Profundidad Emocional indican que, en el caso de los ítems, el promedio del Infit fue de .99 (DT=.27). Para el ajuste global externo, el promedio del Outfit fue de 1.02 (DT=.29). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.00 (DT=.62) y el valor del Outfit de 1.02 (DT= .71). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 8.33% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.36 a 0.74) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

El porcentaje de la varianza explicada por la medida Rasch fue del 47.5%. Asimismo, el primer componente tenía un valor propio de 1.69 lo que corresponde a un 9.9% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos, mediante un Análisis de componentes principales de los residuos (ACPR).

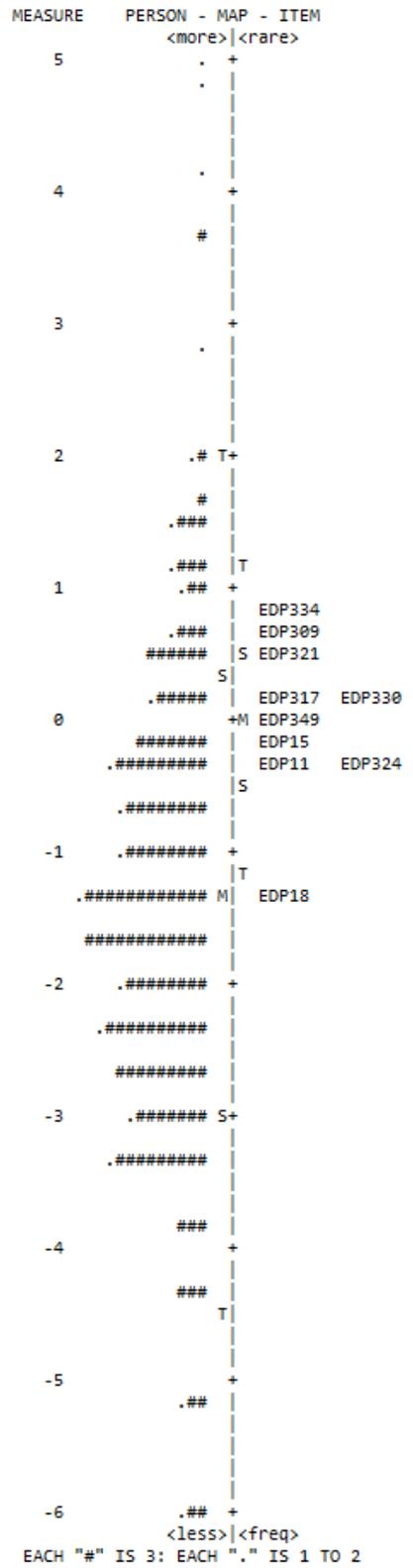
Los ocho ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala Falta de Profundidad Emocional que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.2.4. Escala Falta de Estabilidad Emocional

Figura 12

Gráfico Escala Falta de Estabilidad Emocional – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Falta de Estabilidad Emocional estaba compuesta, originalmente, de 12 ítems. En un primer análisis, se encontró que el ítem 312 podría presentar funcionamiento diferencial (DIF contrast = -0.80, $p < ,001$), siendo de mayor severidad para los hombres de la muestra. En un segundo análisis, el ítem 345 presentó, igualmente, un funcionamiento diferencial (DIF contrast = 1.52, $p < ,001$) de mayor severidad para las mujeres; de esta manera, ambos ítems fueron eliminados.

Ningún otro ítem presentó alteraciones en los valores de infit, outfit, correlación punto media, o afectaciones a la unidimensionalidad. La escala de Falta de Estabilidad Emocional terminó en definitiva con diez ítems en total.

Resultados finales de la escala

Tabla 33

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Falta de Estabilidad Emocional.

	Ítems	Personas
Fiabilidad	.98	.87
Separación	6.48	2.64

Fuente: elaboración propia

En la Tabla 33 se muestran los estadísticos de fiabilidad y separación, que indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos representantes de una falta de estabilidad emocional, por lo que es esperable una consistencia de estas inferencias, una jerarquía en los ítems de la escala. Asimismo, la escala presentó un alfa de Cronbach de .88.

Tabla 34

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Falta de Estabilidad Emocional.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1104	26	-2.88	1.02	1.01	No
2	2140	50	-1.38	.92	.92	-2.84
3	887	21	.26	.92	.92	.28
4	189	4	1.65	1.27	1.33	2.57

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Falta de Estabilidad Emocional cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 35

Estadísticos de los ítems para la escala Falta de Estabilidad Emocional.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM11	-.35	.09	1.14	1.08	.72	1.15	.70	.29
ITEM321	.44	.09	1.21	1.08	1.52	1.22	.62	.33
ITEM324	-.34	.09	1.15	.97	.43	-.34	.66	-.41
ITEM330	.10	.09	1.45	1.08	2.95	1.20	.58	-.02
ITEM349	-.05	.09	.84	.76	-4.53	-4.02	.71	.09
ITEM309	.65	.09	1.05	1.17	3.12	2.46	.70	.00
ITEM317	.20	.09	.80	.95	-.69	-.74	.76	.08
ITEM334	.80	.09	.88	1.07	.77	1.17	.69	.19
ITEM15	-.16	.09	.80	.80	-3.04	-3.19	.75	-.25
ITEM18	-1.29	.08	.67	.97	-.47	-.51	.74	-.24

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Falta de Estabilidad Emocional indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT=.15). Para el ajuste global externo, el promedio del Outfit fue de 1.00 (DT=.23). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de .99 (DT=.64) y el valor del Outfit de 1.00 (DT= .64). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 7.87% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.58 a 0.76) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 49.7%. Asimismo, el primer componente tenía un valor propio de 1.59 lo que corresponde a un 8.0% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

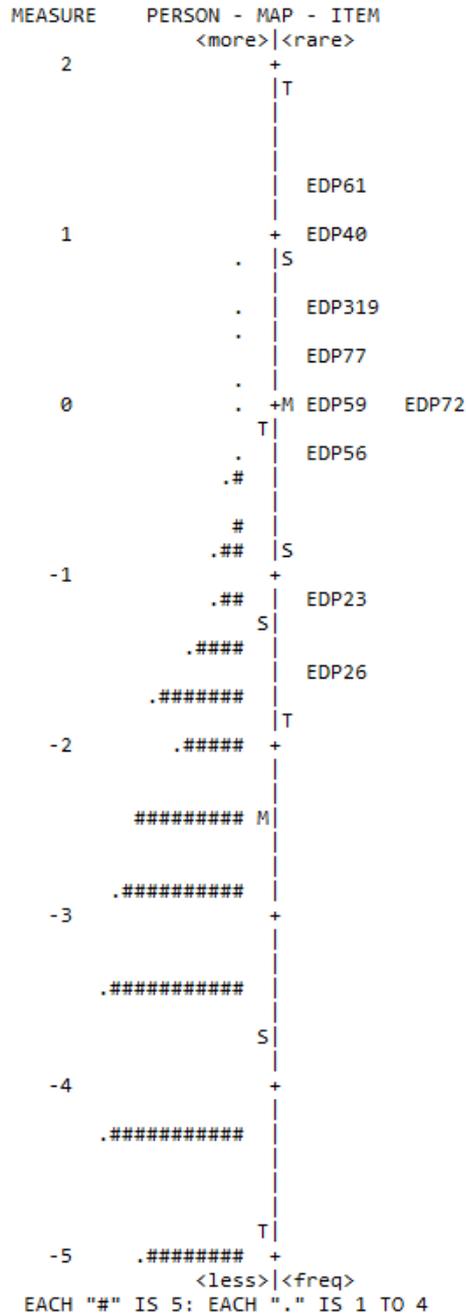
Los diez ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala Falta de Estabilidad Emocional que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.2.5. Escala Falta de Remordimiento

Figura 13

Gráfico Escala Falta de Remordimiento – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Falta de Remordimiento estaba compuesta, originalmente, de diez ítems. En un primer análisis, se encontró que el ítem 44 podría presentar funcionamiento diferencial

(DIF contrast = -0.63, $p < .001$), siendo de mayor severidad para los hombres de la muestra, por lo que fue eliminado.

Ningún otro ítem presentó alteraciones en los valores de infit, outfit, correlación punto media, o afectaciones a la unidimensionalidad. La escala de Falta de Remordimiento terminó en definitiva con nueve ítems en total.

Durante el primer análisis, la confiabilidad de la escala estaba muy baja (.46), tras eliminar el ítem 44 mejoró un poco (.54); sin embargo, aún presentaba una correlación baja. De acuerdo a los criterios de Linacre (2010), se eliminan los casos que presentan valores de Infit y Outfit mayores a 2, 41 en total. En un análisis posterior, la confiabilidad mejoró a .61.

Resultados finales de la escala

Tabla 36

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Falta de Remordimiento.

	Ítems	Personas
Fiabilidad	.99	.61
Separación	8.17	1.24

Fuente: elaboración propia

El alfa de Cronbach es de .76. Con base en este dato y los presentados en la Tabla 36, los estadísticos de fiabilidad y separación, se determina que la escala logra diferenciar a personas con diferentes niveles de rasgos asociados a la falta de remordimiento. Sin embargo, presenta una confiabilidad moderada, pero se esperar consistencia de estas inferencias, y se puede diferenciar una jerarquía en los ítems de la escala.

Tabla 37

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Falta de Remordimiento.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2274	65	-3.19	1.03	1.01	No
2	975	28	-1.72	1.00	.82	-1.81
3	221	6	-.55	1.00	.94	.34
4	49	1	.12	1.25	1.25	1.47

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Falta de Remordimiento cumple con los cuatro criterios propuestos por Linacre (2002) que indican una buena calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 38

Estadísticos de los ítems para la escala Falta de Remordimiento.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM23	-1.21	.08	1.28	1.25	3.38	3.11	.57	.19
ITEM319	.59	.12	.96	.78	-.36	-1.74	.54	.48
ITEM72	-.01	.10	.79	.77	-2.54	-2.42	.62	-.07
ITEM77	.23	.11	.79	.71	-2.47	-2.80	.62	-.14
ITEM26	-1.56	.08	1.08	1.09	1.10	1.24	.60	-.56
ITEM56	-.35	.10	1.08	1.03	.95	.38	.57	.11
ITEM59	-.04	.10	1.07	.93	.84	-.70	.59	.00
ITEM61	1.30	.14	1.11	.88	.95	-.60	.39	.08
ITEM40	1.05	.13	1.00	.85	.05	-.90	.45	.41

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Falta de Remordimiento indican que, en el caso de los ítems, el promedio del Infit fue de 1.02 (DT=.15). Para el ajuste global externo, el promedio del Outfit fue de .92 (DT=.17). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de .99 (DT=.46) y el valor del Outfit de .92 (DT= .51). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Como se mencionó anteriormente, 41 casos (9.49%) de las personas presentaron valores mayores a 2 en Infit u Outfit, lo cual estaba afectando la confiabilidad de la escala y el ajuste de las personas al modelo. Una vez eliminados estos casos, en un siguiente análisis, solo el 4.62% volvió a presentar valores mayores a 2 en Infit u Outfit. De esta forma, parece no haber divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos.

Con el análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.58 a 0.76) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

En lo que respecta al ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR), que mostró un porcentaje de la varianza explicada por la medida Rasch fue del 41%, y el primer componente tenía un valor propio de 1.57 lo que corresponde a un 10.3% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

Los nueve ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

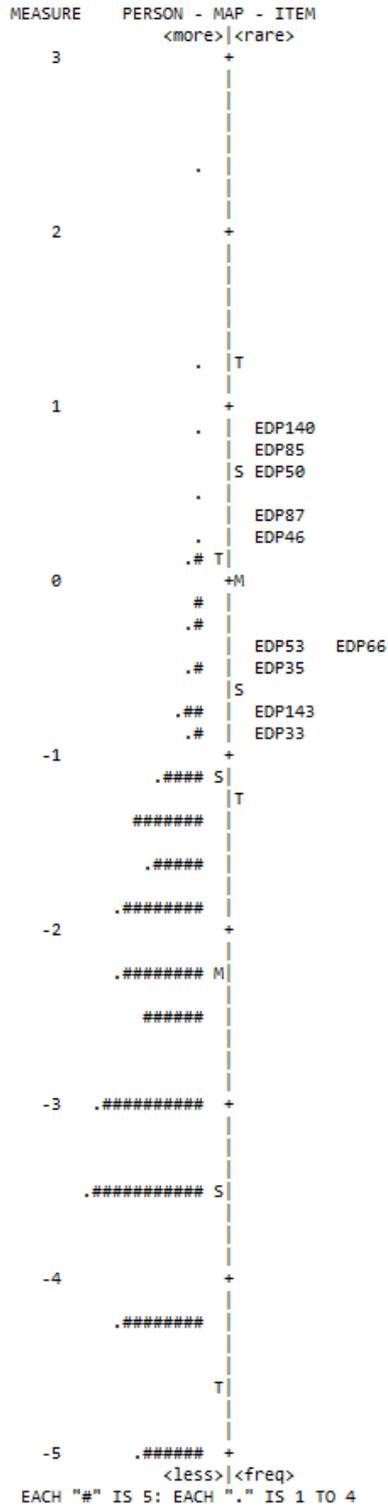
La validez de constructo fue respaldada por todos los ítems en la escala Falta de Remordimiento que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.3. Dominio Dominancia

1.3.1. Escala Antagónico

Figura 14

Gráfico Escala Antagónico – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Antagónico está compuesta por de 10 ítems. Durante la etapa de ajuste y análisis, ningún ítem presentó desajustes o alteraciones en los valores de infit, outfit, correlación punto media, o afectaciones a la unidimensionalidad. Por estas razones no fue necesario eliminar ningún ítem.

Resultados finales de la escala

Tabla 39

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Antagónico.

	Ítems	Personas
Fiabilidad	.98	.70
Separación	6.38	2.64

Fuente: elaboración propia

En la Tabla 39 se muestran los estadísticos de fiabilidad y separación, que indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos antagonistas, por lo que se podría esperar consistencia de estas inferencias, así como una jerarquía en los ítems de la escala. Asimismo, la escala presenta un alfa de Cronbach de = .80, lo que apoya estas conjeturas.

Tabla 40

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Antagónico.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2503	58	-2.95	1.00	.99	No
2	1446	33	-1.67	.96	.89	-1.90
3	316	7	-.71	1.00	.99	.32
4	55	1	.10	1.30	1.53	1.58

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Antagónico cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 41

Estadísticos de los ítems para la escala Antagónico.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM46	.23	.10	1.12	1.10	1.54	1.02	.54	.24
ITEM50	.69	.10	1.23	1.02	2.61	.25	.50	.39
ITEM66	-.35	.09	.78	.84	-3.11	-2.19	.63	-.14
ITEM33	-.88	.08	.89	.90	-1.51	-1.47	.62	-.23
ITEM35	-.44	.08	1.01	1.03	.12	.41	.59	-.18
ITEM53	-.41	.09	.89	.85	-1.47	-2.03	.65	-.30
ITEM85	.72	.11	.91	.81	-1.04	-1.71	.55	.00
ITEM87	.35	.10	.89	.87	-1.34	-1.32	.57	.21
ITEM140	.82	.11	1.15	1.05	1.65	.47	.49	-.02
ITEM143	-.73	.08	1.33	1.36	4.15	4.44	.52	.25

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Antagónico indican que, en el caso de los ítems, el promedio del Infit fue de 1.02 (DT=.17). Para el ajuste global externo, el promedio del Outfit fue de .98 (DT=.16). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT=.50) y el valor del Outfit de .98 (DT= .55). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 6.94% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). Con base en estos datos, se determina que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Adicionalmente, según el análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.49 a 0.63) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

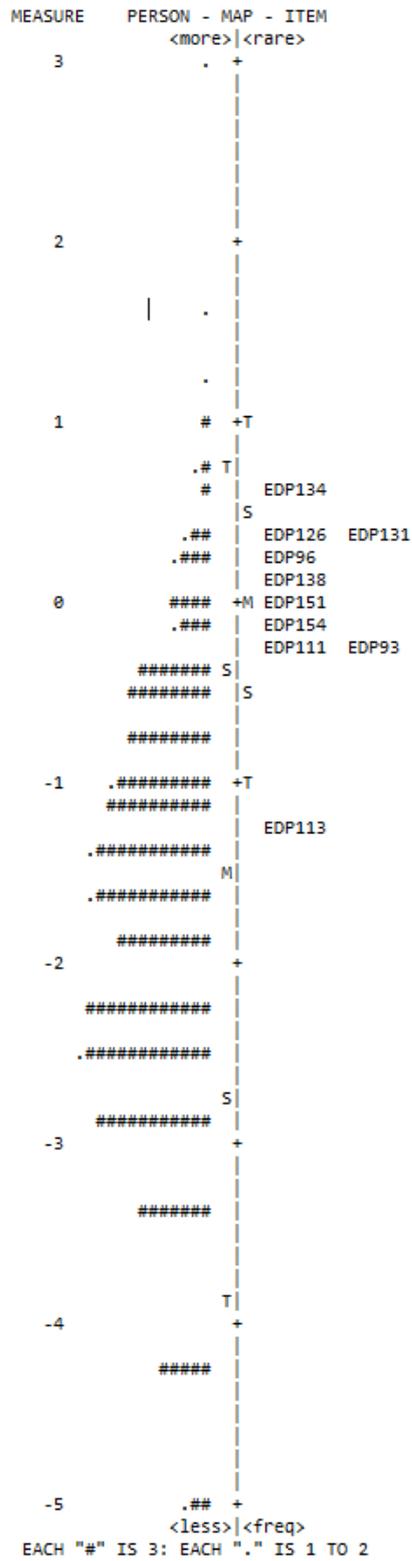
En lo que respecta al supuesto de unidimensionalidad, mediante un Análisis de componentes principales de los residuos (ACPR), el porcentaje de la varianza explicada por la medida Rasch fue del 37.4%, donde el primer componente tenía un valor propio de 1.48 lo que corresponde a un 9.3% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos para las escalas.

Ninguno de los diez ítems permanentes en la escala tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo, presentan un buen funcionamiento diferencial. Asimismo, la validez de constructo fue respaldada por todos los ítems en la escala Antagónico, con un buen ajuste bajo el modelo de Rasch, libres de DIF y correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.3.2. Escala Dominante

Figura 15

Gráfico Escala Dominante – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Dominante constaba de 12 ítems, originalmente, posteriormente a la etapa de análisis y depuración, resultó compuesta de 10. Todos los ítems ajustaron apropiadamente a los valores de infit, outfit, correlación punto media, y no presentaron afectaciones a la unidimensionalidad de la escala. Sin embargo, los ítems 147 y 117 mostraron indicios de funcionamiento diferencial (DIF contrast = 0.645 $p < ,001$ y 0.60, $p < ,001$, respectivamente) (Linacre, 2022), razón por la cual se decidió eliminar de la escala.

Resultados finales de la escala

Tabla 42

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Dominante.

	Ítems	Personas
Fiabilidad	.98	.78
Separación	6.30	1.86

Fuente: elaboración propia

En la Tabla 42 se muestran los estadísticos de fiabilidad y separación, que indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos de dominancia. Junto con un alfa de Cronbach de = .81, se puede determinar una consistencia de estas inferencias y la existencia de una jerarquía en los ítems de la escala.

Tabla 43

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Dominante.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1765	41	-2.38	1.02	1.02	No
2	1750	41	-1.30	.94	.90	-1.87
3	676	16	-.35	.96	.94	.10
4	129	3	.21	1.19	1.23	1.76

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Dominante cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 44

Estadísticos de los ítems para la escala Dominante.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM151	.05	.08	.83	.83	-2.70	-2.43	.66	.36
ITEM93	-.27	.08	1.08	1.09	1.24	1.34	.56	-.15
ITEM126	.43	.08	1.10	1.05	1.40	.59	.56	.46
ITEM131	.34	.08	.94	.83	-.86	-2.20	.66	-.19
ITEM134	.63	.09	.71	.68	-4.52	-4.23	.65	-.27
ITEM154	-.11	.08	1.00	1.05	.09	.68	.56	.00
ITEM111	-.20	.08	.85	.86	-2.31	-2.12	.66	.00
ITEM113	-1.27	.07	1.31	1.34	4.37	4.84	.51	.00
ITEM138	.16	.08	.86	.89	-2.11	-1.52	.58	-.03
ITEM96	.23	.08	1.31	1.27	4.15	3.28	.53	-.19

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Dominante indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT=.19). Para el ajuste global externo, el promedio del Outfit fue de .99 (DT=.20). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.02 (DT=.54) y el valor del Outfit de .99 (DT= .52). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 6.48% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las

personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.51 a 0.66) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 38.9%. Asimismo, el primer componente tenía un valor propio de 1.46 lo que corresponde a un 8.9% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

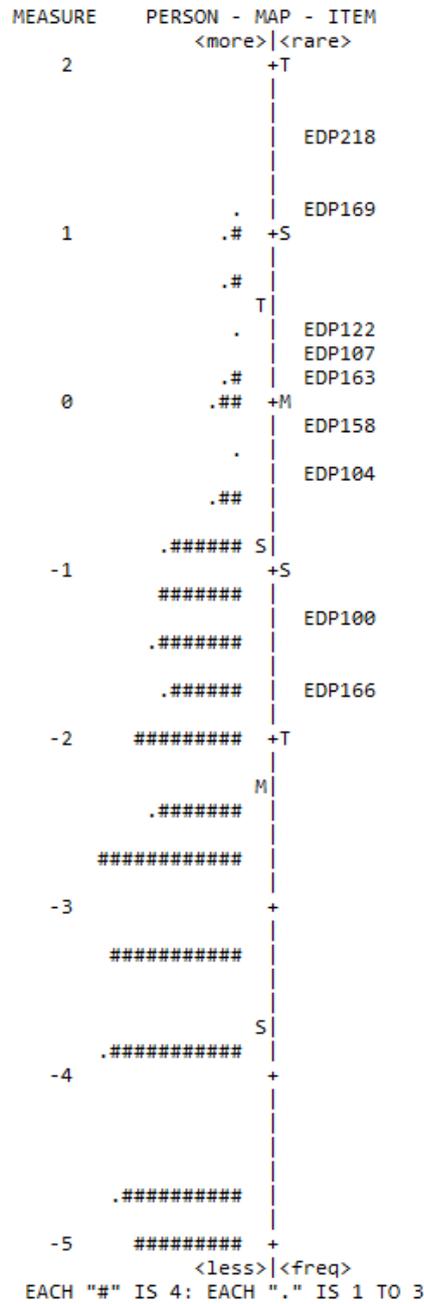
Los diez ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala Dominante que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.3.3. Escala Engañoso

Figura 16

Gráfico Escala Engañoso – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Engañoso estaba conformada inicialmente por 10 ítems, posteriormente a la etapa de análisis y depuración, se decidió eliminar el ítem 206. Este ítem presentó una leve alteración del valor de outfit (1.51) por lo que cumplió el criterio para ser eliminado.

Ningún otro ítem presentó alteraciones a los valores de infit, outfit, correlación punto media, y no presentaron afectaciones a la unidimensionalidad de la escala. La escala terminó compuesta por nueve ítems.

Resultados finales de la escala

Tabla 45

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Engañoso.

	Ítems	Personas
Fiabilidad	.99	.75
Separación	10.16	1.72

Fuente: elaboración propia

En la Tabla 45 se muestran los estadísticos de fiabilidad y separación, estos indican que la escala es eficiente para diferenciar a personas con diferentes niveles de rasgos de personalidad engañosa, por lo que se podría esperar consistencia de estas inferencias. Asimismo, se comprueba que existe una jerarquía en los ítems de la escala. Adicionalmente al valor de fiabilidad, la escala presenta un alfa de Cronbach de = .82.

Tabla 46

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Engañoso.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2096	54	-3.39	1.03	1.01	No
2	1333	34	-1.64	.97	.81	-2.22
3	382	10	-.20	.95	.89	.29
4	77	2	.43	1.37	1.46	1.92

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Engañoso cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 47*Estadísticos de los ítems para la escala Engañoso.*

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM100	-1.25	.08	.96	.95	-.51	-.68	.72	-.26
ITEM104	-.46	.09	1.24	1.23	3.08	2.83	.55	.00
ITEM122	.43	.10	1.07	.89	.90	-1.14	.63	-.02
ITEM158	-.17	.09	.77	.69	-3.29	-4.22	.74	-.30
ITEM163	.20	.10	.87	.91	-1.80	-1.02	.65	.18
ITEM166	-1.71	.08	1.27	1.28	3.62	3.84	.60	-.07
ITEM107	.24	.10	1.09	.98	1.18	-.23	.61	.28
ITEM218	1.59	.12	.95	.72	-.49	-1.73	.51	.39
ITEM169	1.14	.11	.88	.77	-1.47	-1.70	.55	.20

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Engañoso indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT=.16). Para el ajuste global externo, el promedio del Outfit fue de .93 (DT=.20). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.00 (DT=.58) y el valor del Outfit de .93 (DT= .59). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 6.71% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems

obtuvieron valores mayores a .3 (0.51 a 0.74) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Respecto al ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 47.9%. De igual forma, el primer componente tenía un valor propio de 1.83 lo que corresponde a un 10.6% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

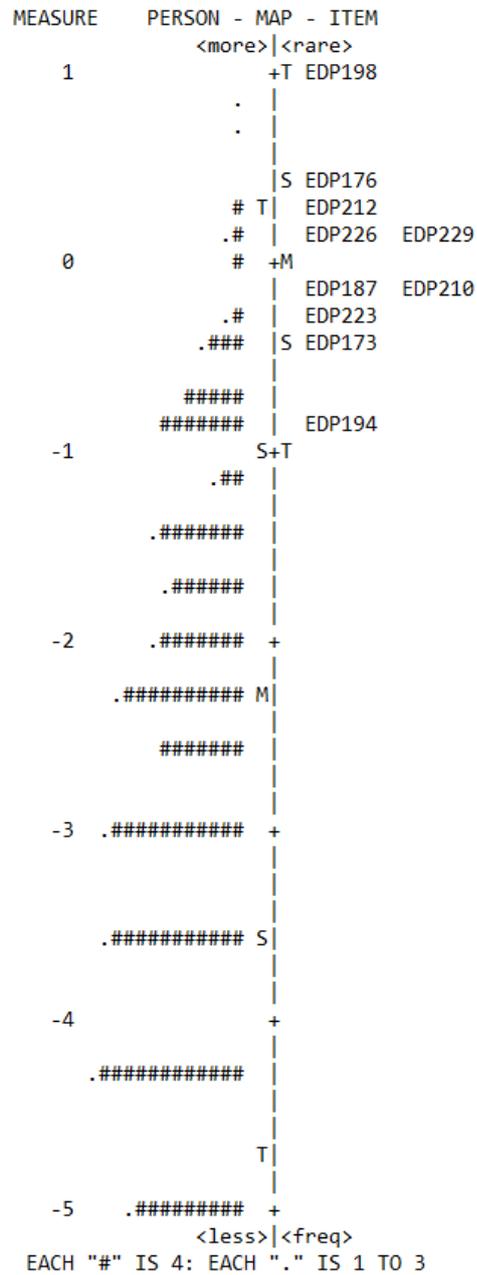
Los nueve ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

Finalmente, la validez de constructo fue respaldada por todos los ítems en la escala Engañoso que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.3.4. Escala Manipulador

Figura 17

Gráfico Escala Manipulador – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Manipulador estaba conformada inicialmente por 11 ítems. Durante la etapa de análisis, se identificó que el ítem 180 presentó indicios de funcionamiento diferencia

de menor severidad en hombres (DIF contrast = 0.62, $p < .001$), por lo que se toma la decisión de eliminarlo de las escala.

El resto de ítems presentó valores apropiados para infit, outfit, correlación punto media, y no hubieron afectaciones a la unidimensionalidad de la escala. La escala final está compuesta por diez ítems.

Resultados finales de la escala

Tabla 48

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Manipulador.

	Ítems	Personas
Fiabilidad	.96	.70
Separación	5.22	1.53

Fuente: elaboración propia

Los estadísticos de fiabilidad y separación, mostrados en la Tabla 48, indican que la escala posee un nivel de sensible moderado para diferenciar a personas con diferentes niveles de rasgos manipuladores, así como consistencia de estas inferencias; además, se comprueba que existe una jerarquía en los ítems de la escala. Esta conclusión es apoyada por un alfa de Cronbach de = .81.

Tabla 49

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Manipulador.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2448	57	-3.00	.97	.97	No
2	1505	35	-1.63	.95	.91	-2.00
3	322	7	-.68	.99	.99	.38
4	45	1	-.70	1.42	1.78	1.63

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Manipulador cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 50

Estadísticos de los ítems para la escala Manipulador.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM210	-.17	.09	1.03	1.05	.36	.59	.59	-.07
ITEM212	.23	.10	.70	.72	-4.13	-3.24	.66	.20
ITEM223	-.32	.09	.89	.84	-1.45	-2.07	.68	.00
ITEM187	-.14	.09	.95	.87	-.58	-1.57	.63	.00
ITEM194	-.93	.08	1.02	1.04	.25	.61	.58	-.50
ITEM198	1.00	.11	.85	.86	-1.80	-1.10	.53	.42
ITEM173	-.43	.09	1.26	1.27	3.26	3.21	.53	.27
ITEM176	.41	.10	1.41	1.31	4.51	2.80	.51	.15
ITEM226	.17	.09	.74	.73	-3.55	-3.21	.65	-.31
ITEM229	.17	.09	1.23	1.31	2.72	3.12	.49	.13

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Manipulador indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT=.22). Para el ajuste global externo, el promedio del Outfit fue de 1.00 (DT=.22). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT=.55) y el valor del Outfit de 1.00 (DT=.59). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 6.01% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las

personas al modelo es aceptable). Por ende, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.49 a 0.68) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 36.3%. Asimismo, el primer componente tenía un valor propio de 1.61 lo que corresponde a un 10.3% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

Los diez ítems en la escala muestran un buen funcionamiento diferencial, no presentan un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo. La validez de constructo fue respaldada por todos los ítems en la escala Manipulador que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.3.5. Escala Insincero

Figura 18

Gráfico Escala Insincero – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Insincero, originalmente conformada por 10 ítems, consta actualmente de 9. Durante el análisis de datos, se identificaron que el ítem 286 y 273 presentaron indicios

de funcionamiento diferencia, el 286 de menor severidad en hombres (DIF contrast = 0.67, $p < .001$), y el 273 por menor severidad las mujeres de la muestra (DIF contrast = -0.66, $p < .001$). Se procede a eliminar el ítem 286; sin embargo, con base un análisis de contenido y para preservar la fiabilidad de la escala, se decide mantener el ítem 273. El resto de ítems presentó valores apropiados para infit, outfit, correlación punto media, y no hubieron afectaciones a la unidimensionalidad de la escala.

Resultados finales de la escala

Tabla 51

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Insincero.

	Ítems	Personas
Fiabilidad	.99	.77
Separación	8.89	1.81

Fuente: elaboración propia

En la Tabla 51, los estadísticos de fiabilidad y separación indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos de insinceridad y que se podría esperar consistencia de estas inferencias, además de una jerarquía en los ítems de la escala. Asimismo, la escala presenta un alfa de Cronbach de = .79.

Tabla 52

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Insincero.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1622	42	-2.52	.95	.96	No
2	1518	39	-1.20	.99	.94	-1.88
3	617	16	-.23	.96	.96	.15
4	131	3	.29	1.29	1.38	1.73

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Insincero cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 53

Estadísticos de los ítems para la escala Insincero.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM233	-.25	.08	1.01	1.04	.22	.58	.58	-.03
ITEM236	.45	.08	1.19	1.09	2.50	1.10	.58	-.16
ITEM182	.44	.08	1.12	1.17	1.66	2.03	.46	.17
ITEM282	.33	.08	.86	.79	-1.97	-2.87	.65	.38
ITEM289	.09	.08	.92	.94	-1.22	-.77	.59	.15
ITEM240	.10	.08	1.13	1.17	1.84	2.21	.57	.32
ITEM269	1.16	.09	1.01	.89	.13	-1.02	.55	.46
ITEM273	-1.11	.07	1.04	.05	.59	.74	.60	-.60
ITEM256	-1.22	.07	.87	.85	-2.04	-2.37	.68	-.32

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Insincero indican que, en el caso de los ítems, el promedio del Infit fue de 1.02 (DT=.11). Para el ajuste global externo, el promedio del Outfit fue de 1.00 (DT=.13). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.00 (DT=.60) y el valor del Outfit de 1.00 (DT=.62). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos, solamente 8.56% de las personas presentaron valores mayores a 2 en Infit u Outfit, lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable. Posterior al análisis de los valores de

correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.46 a 0.68) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 42.7%. Asimismo, el primer componente tenía un valor propio de 1.82 lo que corresponde a un 11.6% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

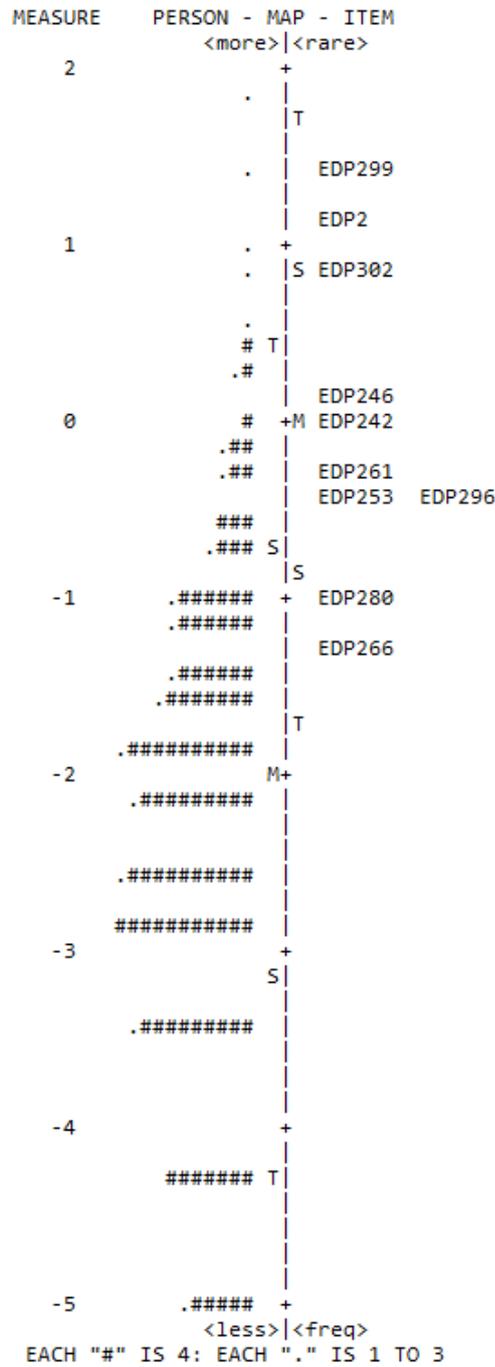
Ninguno de los ítems, aparte del ya mencionado caso del ítem 273, tienen un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo; de esta manera, la escala muestra un buen funcionamiento diferencial.

En conclusión, la validez de constructo fue respaldada por los nueve ítems en la escala Insincero que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.3.6. Escala Charlatán

Figura 19

Gráfico Escala Charlatán – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Charlatán está conformada por de 10 ítems. Durante la etapa de ajuste y análisis, ningún ítem presentó desajustes o alteraciones en los valores de infit, outfit,

correlación punto media, o afectaciones a la unidimensionalidad. Con base en lo anterior, no fue necesario eliminar ningún ítem.

Resultados finales de la escala

Tabla 54

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Charlatán.

	Ítems	Personas
Fiabilidad	.99	.74
Separación	9.63	1.67

Fuente: elaboración propia

En la Tabla 54 se muestran los estadísticos de fiabilidad y separación; además de un alfa de Cronbach de = .81, indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos de charlatanería. Adicionalmente, presenta valores adecuados para una consistencia de estas inferencias, así como una jerarquía en los ítems de la escala.

Tabla 55

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Charlatán.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2290	53	-2.82	.92	.95	No
2	1424	33	-1.40	.97	.88	-1.73
3	497	12	-.34	.95	.98	.13
4	109	3	.10	1.43	1.91	1.60

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala Charlatán cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 56*Estadísticos de los ítems para la escala Charlatán.*

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM261	-.30	.08	1.02	1.06	.35	.73	.58	.09
ITEM280	-1.04	.07	1.03	1.02	.39	.30	.64	-.05
ITEM242	-.02	.08	.90	.90	-1.38	-1.16	.61	-.16
ITEM246	.10	.08	1.14	1.05	1.88	-.56	.60	.00
ITEM266	-1.32	.07	1.03	1.04	.43	.68	.59	-.24
ITEM296	-.37	.08	.76	.74	-3.64	-3.70	.67	.06
ITEM299	1.37	.11	1.13	.94	1.33	-.33	.43	.19
ITEM302	.83	.10	1.27	1.34	2.96	2.67	.47	.32
ITEM253	-.45	.08	1.02	1.08	.29	1.10	.61	.07
ITEM2	1.20	.11	.90	.98	-1.15	-.13	.48	-.03

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala Charlatán indican que, en el caso de los ítems, el promedio del Infit fue de 1.02 (DT=.14). Para el ajuste global externo, el promedio del Outfit fue de 1.02 (DT=.14). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT=.59) y el valor del Outfit de 1.02 (DT=.76). De manera que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente.

Por otro lado, solamente 9.25% de las personas presentaron valores mayores a 2 en Infit u Outfit, por lo que se determina que el ajuste de las personas al modelo es aceptable. De esta manera, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Asimismo, de acuerdo con el análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores

mayores a .3 (0.43 a 0.67) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

El porcentaje de la varianza explicada por la medida Rasch fue del 44.4%, y el primer componente con un valor propio de 1.64, fue de 9.2% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos, según un Análisis de componentes principales de los residuos (ACPR).

De los diez ítems que se mantuvieron en la escala, todos muestran un buen funcionamiento diferencial, y ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala Charlatán que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.4. Dominio de Apego

1.4.1. Escala Desapegado

Figura 20

Gráfico de Escala Desapegado – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Desapegado originalmente constaba de 11 ítems, durante el proceso de ajuste y revisión de ítems, se tomó la decisión de conservar la totalidad de los mismos. Esto se debe a que no se encontró desajustes severos en el infit y tampoco en el outfit. Por otro

lado, los ítems no presentaron una correlación punto medida baja, por lo cual tampoco se descartaron ítems bajo este criterio. Asimismo, no se encontraron ítems que afectaran la unidimensionalidad. Finalmente, se evaluó si algún ítem podría estar presentando funcionamiento diferencial, siguiendo con los criterios de Linacre (2022), pero nuevamente no fue el caso, los ítems mostraron puntajes de DIF Contrast entre los rangos de .00 y .40. Tomando en cuenta todo lo mencionado anteriormente, se decidió mantener la Escala de Desapegado con su totalidad de 11 ítems.

Resultados finales de la escala

Tabla 57

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Desapegado

	Ítems	Personas
Fiabilidad	.98	.80
Separación	7.38	2.02

Fuente: elaboración propia

Tal y como se observa en la Tabla 57, los estadísticos de fiabilidad y separación indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos de desapego y que podríamos esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. Adicionalmente, el puntaje de Alpha de Cronbach para esta escala fue de .80, mismo que se puede considerar como bastante elevado (Taber, 2017), y que reafirma la habilidad de esta escala para evaluar el constructo.

Tabla 58

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Rasgos Impulsivos.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1618	34	-1.83	1.00	1.00	No
2	1902	40	-.86	.90	.90	-1.48
3	915	19	-.07	.91	.93	.23
4	317	7	.38	1.16	1.32	1.25

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la Escala Desapegado, cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 59

Estadísticos de los ítems para la escala Rasgos Impulsivos

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM9	.15	.07	1.05	1.05	.71	.66	.57	-.10
ITEM306	.35	.07	.86	.85	-2.14	-2.05	.61	.00
ITEM342	-.69	.06	1.01	1.00	.16	.03	.61	.00
ITEM346	.31	.07	1.07	1.03	1.08	.42	.54	-.40
ITEM12	-.64	.06	1.30	1.41	4.46	5.63	.42	-.16
ITEM322	.23	.07	.85	.87	-2.26	-1.87	.59	-.18
ITEM325	-.48	.07	.66	.65	-6.08	-6.02	.70	-.08
ITEM328	.08	.07	1.33	1.31	4.44	3.92	.56	.16
ITEM331	.86	.08	1.15	1.26	2.02	2.78	.44	.49
ITEM350	.54	.07	.93	1.00	-1.06	.03	.50	.35
ITEM313	-.71	.06	.82	.82	-3.15	-2.92	.65	.03

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la Escala de Desapegado indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT= .20). Para el ajuste global externo, el promedio del Outfit fue de 1.02 (DT=.22). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.02 (DT= .54) y el valor del Outfit de 1.02 (DT= .58). De esta forma, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente. Por otro lado, solamente el 6% de las personas

presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.42 a 0.70) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 39.4%. Asimismo, el primer componente tenía un valor propio de 1.83 lo que corresponde a un 10.1 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

Los once ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la Escala de Desapegado que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.4.2. Escala de No Adquiere Compromisos

Figura 21

Gráfico Escala No Adquiere Compromisos – Distribución casos e ítems



Resultados iniciales de la escala

La escala de No Adquiere Compromisos, originalmente constaba de 10 ítems, durante el proceso de ajuste se tomó la decisión de eliminar un ítem. El ítem 24 mostró un desajuste

severo tanto en el infit, como en el outfit (infit=1.58, outfit = 1.56). Adicionalmente, se tomó la decisión de eliminar parte de la muestra al analizar esta escala, esto se decidió al seguir las recomendaciones de Linacre (2010) sobre cuando eliminar y dejar de eliminar ítems o personas que muestren un severo desajuste con la escala. Bajo este criterio, se eliminó un total de 40 personas, mejorando los estadísticos de fiabilidad con un aumento del .7 en el caso de los ítems y .6 en las personas, mientras que en separación la mejoría fue de .20 en el caso de los ítems, y de .64 en el caso de las personas. Estos resultados son un claro indicador de que parte de los ajustes necesarios para una mejor interpretación y validación de esta escala, está en conseguir una muestra más amplia y heterogénea.

Por otro lado, ningún ítem presentó una correlación punto medida baja. A su vez, la unidimensionalidad no se vio afectada por ninguno de los ítems. Finalmente, siguiendo con los criterios de Linacre (2022), se realizaron análisis para determinar si algún ítem estaba presentando funcionamiento diferencial, pero este no fue el caso. La escala de No Adquiere Compromisos quedó en definitiva con nueve ítems en total.

Resultados finales de la escala

Tabla 60

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de No Adquiere Compromisos

	Ítems	Personas
Fiabilidad	.99	.66
Separación	9.48	1.39

Fuente: elaboración propia

Tal y como se observa en la Tabla 60, los estadísticos de fiabilidad y separación indican que hay espacio para mejoras en la escala, para que la misma sea más sensible para diferenciar a personas con diferentes niveles de rasgos de “no adquirir compromisos”. Estas mejoras pueden llevar a una mayor consistencia de estas inferencias, y a su vez, se lograría mostrar con mayor claridad, que existe una jerarquía en los ítems de la escala. Adicionalmente, el valor del Alpha de Cronbach para esta escala fue de .75, un puntaje que se puede considerar como alto (Taber, 2017) y reforzando el hecho de que esta tabla muestra

potencial y datos aceptables, pero que requieren una mejora para que su habilidad de evaluación del constructo se pueda considerar verdaderamente sólida.

Tabla 61

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de No Adquiere Compromisos.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2103	60	-3.47	1.03	1.02	No
2	1142	32	-1.88	1.03	.92	-2.17
3	246	7	-.53	.94	.86	.33
4	37	1	.42	1.03	1.00	1.84

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala de No Adquiere Compromisos, se ajusta de manera adecuada a los cuatro criterios propuestos por Linacre (2022) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 62*Estadísticos de los ítems para la escala de No Adquiere Compromisos*

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM31	.77	.12	1.05	1.18	.54	1.36	.45	-.16
4	1.76	.15	.92	.56	-.64	-2.31	.46	.30
ITEM33	-.48	.09	.97	1.10	-.40	1.21	.48	-.27
5	.56	.11	.85	.73	-1.80	-2.46	.56	.00
ITEM16	-.83	.09	.99	.95	-.13	-.67	.63	-.13
ITEM19	-.06	.10	.91	.82	-1.17	-2.00	.63	.40
ITEM21	-1.54	.08	1.01	1.00	.20	.05	.68	.32
ITEM70	-1.20	.09	1.21	1.21	2.65	2.69	.58	-.05
ITEM73	1.02	.13	1.06	.92	.65	-.47	.44	-.47
ITEM75								
ITEM57								

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de No Adquiere Compromisos, indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT= .10). Para el ajuste global externo, el promedio del Outfit fue de .94 (DT=.20). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.00 (DT= .42) y el valor del Outfit de .94 (DT= .50). De esta forma, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente. Por otro lado, solamente 4.59% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos confirman que no existen discrepancias significativas entre las predicciones

del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.44 a 0.68) por lo tanto, se confirma que los mismos cumplen con el criterio de validez de contenido requerido.

Para determinar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 43.1%. Asimismo, el primer componente tenía un valor propio de 1.54 lo que corresponde a un 9.8 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

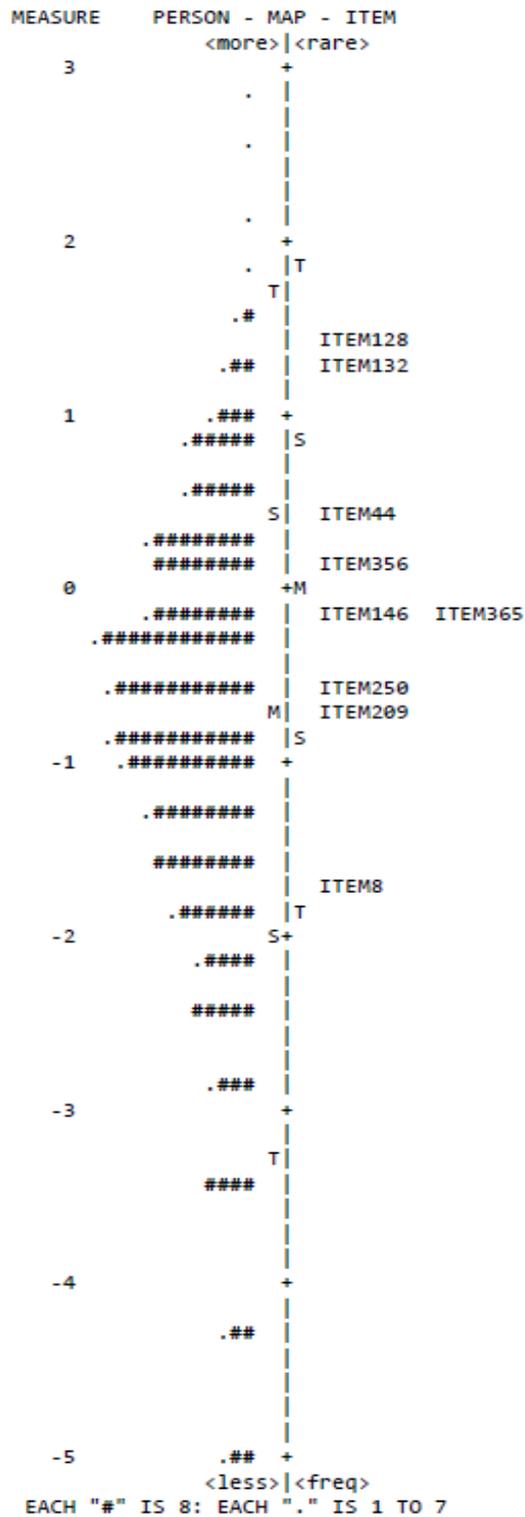
Los nueve ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de No Adquiere Compromisos, que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.4. 3. Escala de Sin Empatía

Figura 22

Gráfico Escala Sin Empatía – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Sin Empatía originalmente constaba de 10 ítems, durante el proceso de ajuste se eliminó un ítem. Inicialmente se realizaron análisis para determinar si habían ítems que desajustaron severamente el infit y/o el outfit, pero los resultados en este apartado fueron positivos, por lo cual no se eliminó ningún ítem bajo este criterio. Por otro lado, los ítems tampoco presentaron una correlación punto medida baja, por lo cual este criterio tampoco fue determinante para la decisión de eliminar uno de los ítems. La unidimensionalidad tampoco estaba siendo afectada por los ítems. Finalmente, se hizo un análisis siguiendo los criterios de Linacre (2022), y en este caso detectó un ítem que podría estar presentando funcionamiento diferencial, por lo que se tomó la decisión de eliminar el ítem 79 (DIF contrast = 0.54, $p < .001$) siendo de menor severidad para los hombres de la muestra. La escala de Sin Empatía, quedó en definitiva con nueve ítems en total.

Resultados finales de la escala

Tabla 63

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Sin Empatía

	Ítems	Personas
Fiabilidad	.97	.66
Separación	6.04	1.40

Fuente: elaboración propia

En la tabla anterior podemos observar, que los estadísticos de fiabilidad y separación indican que es posible una mejoría en la escala para que esta sea más sensible para diferenciar a personas con diferentes niveles de falta de empatía, y que podríamos esperar aún mayor consistencia de estas inferencias con algunos ajustes a los ítems, o bien con una muestra más grande y heterogenea. De igual forma, estos cambios podrían comprobar con mayor claridad la existencia de una jerarquía en los ítems de la escala. Adicionalmente, el valor del Alpha de Cronbach fue de .76, mismo que se puede considerar alto según Taber (2017), pero que, a su vez, refuerza lo antes expuesto, al ser necesario mejorías para fortalecer la confiabilidad de la escala.

Tabla 64

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Sin Empatía.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2109	54	-2.53	1.01	1.00	No
2	1370	35	-1.52	.94	.94	-1.70
3	336	9	-.50	.92	.92	.33
4	73	2	-.10	1.33	1.53	1.37

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala de Sin Empatía cumple con los cuatro criterios propuestos por Linacre (2002), mismos que confirman la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 65

Estadísticos de los ítems para la escala Sin Empatía

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfi	Infit	Outfi		
ITEM62	.32	.09	1.39	1.45	4.45	4.37	.46	.26
ITEM41	-.82	.07	.68	.70	-5.13	-4.67	.66	.00
ITEM45	.63	.10	.88	.80	-1.44	-2.12	.55	.00
ITEM47	.49	.09	.89	.92	-1.40	-.84	.49	-.19
ITEM67	-.11	.08	1.21	1.18	2.66	2.21	.53	-.07
ITEM28	-.37	.08	1.10	1.12	1.40	1.61	.54	-.42
ITEM36	.30	.09	.96	.93	-.53	-.71	.56	.05
ITEM54	.39	.09	.84	.84	-2.07	-1.78	.57	.35
ITEM82	-.83	.07	1.10	1.10	1.45	1.34	.59	.12

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Sin Empatía indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT= .20). Para el ajuste global externo, el promedio del Outfit fue de 1.01 (DT=.22). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT= .52) y el valor del Outfit de 1.01 (DT= .54). De esta forma, ninguno de los ítems muestra indicios de desajuste (Infit u Outfit > 1.5), por lo que se comprueba que los valores de ajuste de todos los ítems se encuentran en el rango esperado para obtener una medición adecuada. Por otro lado, solamente 6.48% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos muestran que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.46 a 0.66) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 36.1%. Asimismo, el primer componente tenía un valor propio de 1.74 lo que corresponde a un 12.4 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

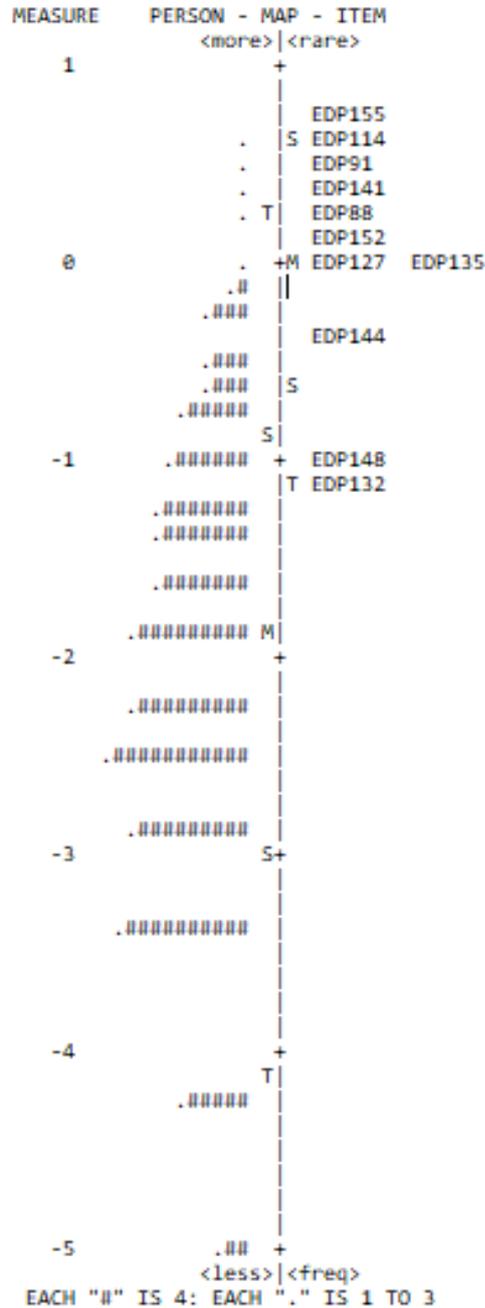
Los nueve ítems que conforman la escala muestran un buen funcionamiento diferencial. Ninguno de ellos muestra un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Sin Empatía, que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.4.4. Escala de Indiferente

Figura 23

Gráfico Escala de Indiferente – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Indiferente originalmente constaba de 11 ítems, durante el proceso de análisis, se tomo la decisión de no eliminar ningún ítem. Esto se debe a que no se encontró que los ítems desajustaran severamente en el infit y/o el outfit. Por otro lado, los ítems

tampoco mostraron una correlación punto medida baja. De igual forma, la unidimensionalidad no se vio afectada por este grupo de ítems. Finalmente, al seguir los criterios de Linacre (2022), no se detectaron ítems que presentaran un funcionamiento. Con lo cual, la escala de Indiferente quedó en definitiva con los once ítems en total.

Resultados finales de la escala

Tabla 66

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Indiferente

	Ítems	Personas
Fiabilidad	.98	.72
Separación	6.89	1.59

Fuente: elaboración propia

Tal y como se observa en la Tabla 66, los estadísticos de fiabilidad y separación indican que la escala muestra potencial para ser sensible y diferenciar a personas con distintos niveles de rasgos de indiferencia y que se podría esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. Finalmente, el valor del Alpha de Cronbach fue de .76, mismo que se puede considerar alto (Taber, 2017) y que respalda que el hecho de que esta escala, presenta datos sólidos para ser considerada como hábil para evaluar este constructo, pero que, a su vez, muestra potencial para ser mejorada y reforzar su confiabilidad.

Tabla 67

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Indiferente.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2476	52	-2.60	.95	.96	No
2	1768	37	-1.39	.93	.84	-1.72
3	418	9	-.64	1.04	1.08	.41
4	90	2	-.36	1.36	1.71	1.31

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala de Indiferente cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 68

Estadísticos de los ítems para la escala de Indiferente

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM88	.20	.09	1.05	1.02	.71	.31	.48	-.05
ITEM91	.48	.09	.85	.79	-2.04	-2.32	.56	.08
ITEM141	.41	.09	.85	.82	-2.03	-2.01	.56	-.09
ITEM144	-.43	.08	1.10	1.18	1.44	2.34	.50	.15
ITEM148	-1.02	.07	.68	.75	-5.44	-4.08	.56	-.43
ITEM152	.16	.09	1.13	1.01	1.65	.16	.56	-.41
ITEM127	.02	.08	1.03	.97	.40	-.32	.54	-.12
ITEM132	-1.12	.07	1.27	1.28	3.71	3.86	.52	.48
ITEM135	-.06	.08	.81	.76	-2.71	-3.28	.60	-.07
ITEM155	.69	.10	.21	1.20	2.48	1.87	.44	.43
ITEM114	.67	.10	1.40	1.32	4.41	2.87	.42	.09

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Imprudente indican que, en el caso de los ítems, el promedio del Infit fue de 1.03 (DT= .21). Para el ajuste global externo, el promedio del Outfit fue de 1.01 (DT=.20). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.02 (DT= .57) y el valor del Outfit de 1.01 (DT= .61). De esta forma, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente. Por otro lado, solamente 7.87% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que

permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.42 a 0.60) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 34.4%. Asimismo, el primer componente tenía un valor propio de 1.67 lo que corresponde a un 10% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

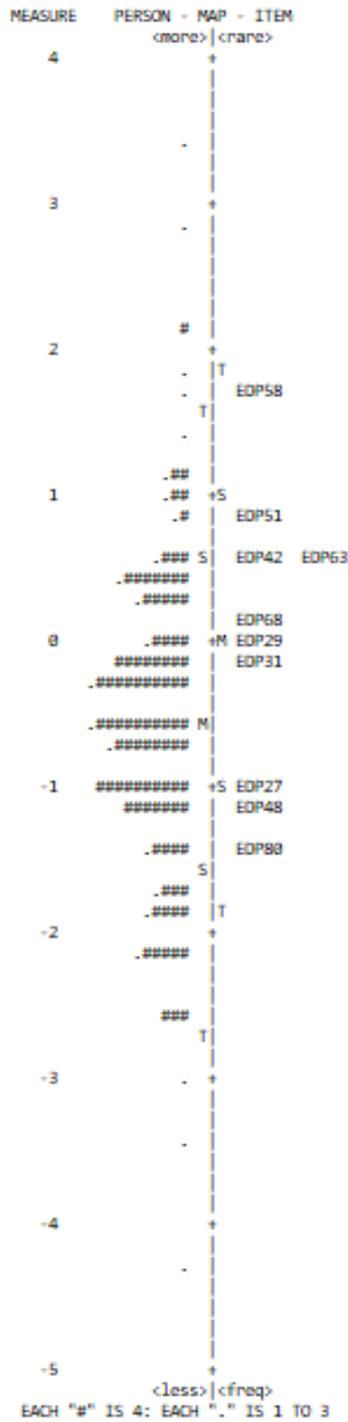
Los once ítems de la escala muestran un buen funcionamiento diferencial, ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo. La validez de constructo fue respaldada por todos los ítems en la escala de Indiferente que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.5. Dominio Cognitivo

1.5.1. Escala de Suspica

Figura 24

Gráfico Escala Suspica – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Suspicaaz originalmente constaba de 10 ítems, durante el proceso de revisión y ajuste se llegó a la conclusión de no eliminar ítems. Esto se debe a que al realizar los diferentes análisis, ninguno de los ítems se salía de los criterios establecidos. Tal es el caso infit y el outfit, donde ningún ítem desajustó severamente. Por otro lado, los ítems tampoco arrojaron una correlación punto medida baja. A su vez, la unidimensionalidad no se vio afectada. Finalmente, al seguir los criterios de Linacre (2022), no se detectaron dos ítems que presentaran un funcionamiento diferencial, por lo cual, se decide conservar la totalidad de los 10 ítems en la escala de Suspicaaz.

Resultados finales de la escala

Tabla 69

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Suspicaaz

	Ítems	Personas
Fiabilidad	.99	.80
Separación	13.30	2.02

Fuente: elaboración propia

Tal y como se observa en la Tabla 69, los estadísticos de fiabilidad y separación indican que la escala es sensible para diferenciar a personas con diferentes niveles de suspicacia y que podríamos esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. Adicionalmente, el Alfa de Cronbach obtenido fue de .78, mismo que se puede considerar como bastante alto según Taber (2017), y que respalda la confiabilidad de esta escala para detectar el constructo de suspicaaz.

Tabla 70

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala Suspica.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1086	25	-1.91	.97	.98	No
2	1648	38	-.71	1.01	.99	-1.70
3	1092	25	.29	.91	.86	.16
4	496	11	1.19	1.15	1.17	1.54

Fuente: elaboración propia

Los datos obtenidos de las categorías de respuesta muestran que la escala de Suspica cumple con los cuatro criterios propuestos por Linacre (2002) que corroboran la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 71

Estadísticos de los ítems para la escala de Suspica

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM27	-1.04	.07	.78	.79	-3.87	-3.42	.59	-.26
ITEM58	1.74	.09	.98	.90	-.31	-1.12	.54	.11
ITEM63	.50	.07	.76	.75	-3.97	-4.02	.61	-.07
ITEM80	-1.45	.07	1.06	1.10	.89	1.49	.48	-.21
ITEM42	.59	.07	.99	.93	-.07	-.99	.59	-.12
ITEM48	-1.21	.07	1.02	1.04	.29	.56	.55	-.32
ITEM51	.88	.07	.95	.89	-.67	-1.52	.68	.26
ITEM68	.20	.07	1.12	1.12	1.78	1.73	.62	.45
ITEM29	-.03	.07	1.33	1.32	4.64	4.52	.54	.12
ITEM31	-.17	.07	1.09	1.13	1.35	1.94	.59	.14

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de SuspicaZ indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT= .16). Para el ajuste global externo, el promedio del Outfit fue de 1.00 (DT=.17). En cuanto a las personas, estos indicadores mostraron que el valor del ajuste interno obtenido fue de 1.01 (DT= .60) y el valor del Outfit de 1.00 (DT= .59). De esta forma, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se demuestra que los valores de ajuste de todos los ítems se incluyen en el rango deseado para una medición adecuada. Por otro lado, solamente 7.63% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.52 a 0.68) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 48.7%. Asimismo, el primer componente tenía un valor propio de 1.93 lo que corresponde a un 9.9 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

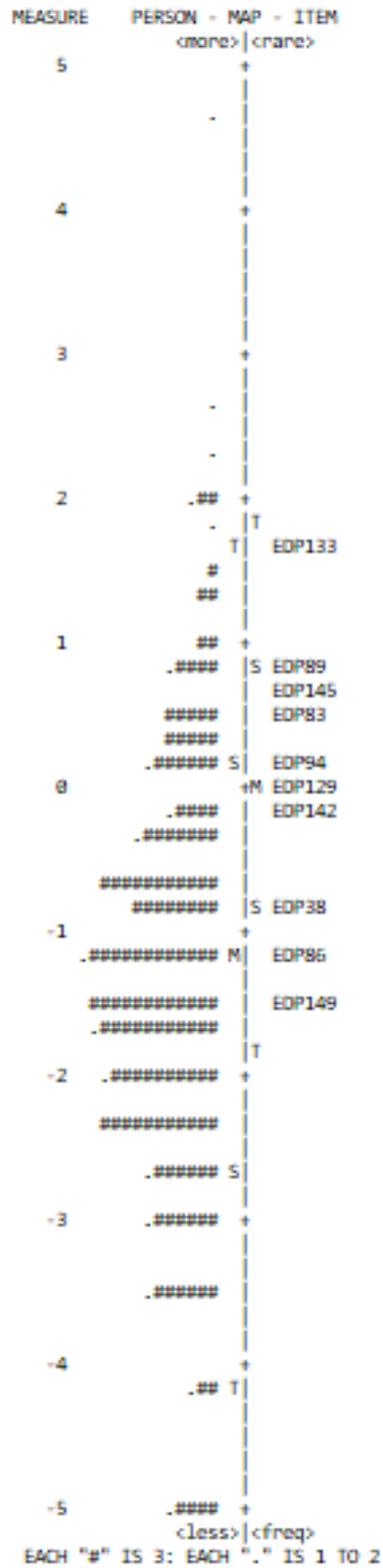
Los diez ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de SuspicaZ que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.5.2. Escala de Falta de Concentración

Figura 25

Gráfico Escala de Falta de Concentración – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Falta de Concentración originalmente constaba de 10 ítems, durante el proceso de ajuste se tomó la decisión de no eliminar ninguno. Esta decisión se debe a que ningún ítem desajustó severamente el infit y, ni el outfit. Por otro lado, los ítems no presentaron una correlación punto medida baja, por lo cual no hubo descartes bajo este criterio. A su vez, ningún ítem afectaba la unidimensionalidad de la escala de Falta de Concentración. Finalmente, se siguieron los criterios de Linacre (2022), pero en los resultados no se detectaron ítems que podrían presentar un funcionamiento diferencial, con lo cual, la escala de Falta de Concentración quedó en definitiva con diez ítems en total.

Resultados finales de la escala

Tabla 72

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Falta de Concentración

	Ítems	Personas
Fiabilidad	.99	.85
Separación	10.68	2.40

Fuente: elaboración propia

Tal y como se observa en la Tabla 72, los estadísticos de fiabilidad y separación confirman que la escala es sensible para diferenciar a personas con diferentes niveles de falta de concentración y que se puede esperar consistencia de estas inferencias, así como que se corrobora que hay una jerarquía en los ítems de la escala. Asimismo, el Alfa de Cronbach obtenido fue de .86, puntaje que puede ser considerado como confiable (Taber, 2017) y que respalda la solidez de la escala para evidenciar rasgos de falta de concentración.

Tabla 73

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Falta de Concentración

Categoría	Muestra	Frecuencia	Promedio	InFit	OutFit	Paso
		(%)				
1	1282	30	-2.78	1.00	1.00	No
2	1939	45	-1.22	.95	.92	-2.41
3	874	20	.25	.89	.87	.26
4	225	5	1.21	1.29	1.34	2.16

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala de Falta de Concentración cumple a cabalidad con los criterios planteados por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 74

Estadísticos de los ítems para la escala de Falta de Concentración

Ítem	Nivel de	Error	MNSQ		ZTSD		Correlacione	DIF
			severidad	estándar	Infit	Outfit		
	(en logits)						medida	
ITEM83	.42	.08	.94	.95	-.95	-.64	.66	-.12
ITEM86	-1.17	.08	1.22	1.19	3.13	2.67	.68	.00
ITEM89	.77	.09	.91	.87	-1.38	-1.75	.69	.28
ITEM38	-.79	.08	.73	.74	-4.46	-4.17	.68	-.19
ITEM142	-.22	.08	.77	.77	-3.73	-3.58	.73	-.34
ITEM145	.66	.09	.96	.93	-.58	-.89	.64	.33
ITEM149	-1.42	.08	1.21	1.29	3.09	3.92	.52	.00
ITEM94	.20	.08	1.26	1.26	3.54	3.40	.58	.00
ITEM129	-.06	.08	.83	.79	-2.65	-3.29	.75	-.10
ITEM133	1.61	.10	.1.2	1.11	3.51	1.02	.54	.26

7

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Falta de Concentración indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT= .20). Para el ajuste global externo, el promedio del Outfit fue de 0.99 (DT=.20). En cuanto a los resultados de las personas, el valor del ajuste interno obtenido fue de 1.00 (DT= .61) y el valor del Outfit de 0.99 (DT= .60). De esta forma, no existen ítems presenten índices de ajuste inadecuados (Infit u Outfit $>$ 1.5), por lo que se confirma que los valores de ajuste de todos los ítems se incluyen en el rango necesario para medir adecuadamente. Por otro lado, solamente 7.87% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, se puede interpretar que estos datos señalan que no hay diferencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se determinó que todos los ítems obtuvieron valores mayores a .3 (0.52 a 0.75) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para determinar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 50.6%. Asimismo, el primer componente tenía un valor propio de 1.75 lo que corresponde a un 8.7 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad esperados.

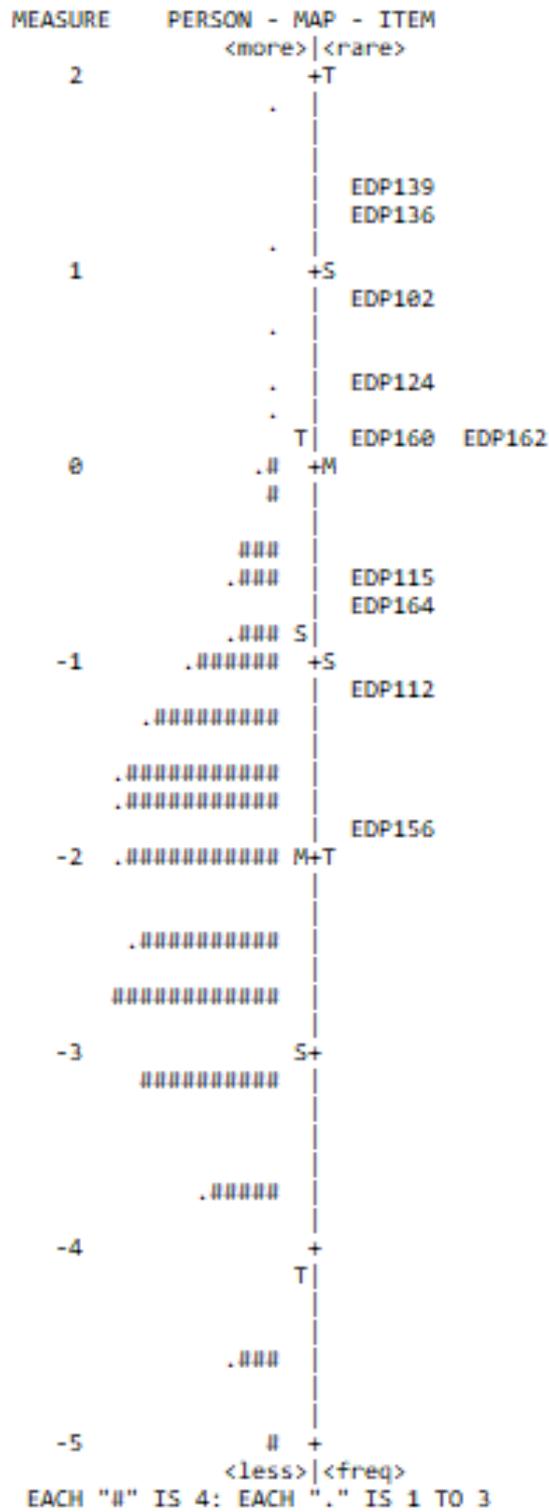
Los diez ítems que se mantuvieron en la escala muestran un alto funcionamiento diferencial. Ninguno muestra un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Falta de Concentración que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.5.3. Escala de Intolerante

Figura 26

Gráfico Escala de Intolerante – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Intolerante originalmente constaba de 12 ítems, durante el proceso de ajuste se eliminaron los siguientes dos ítems. El ítem 98 desajustó severamente el infit y en el outfit (infit=1.97, outfit = 1.73), adicionalmente el ítem 120 también mostró un desajuste severo en ambos rubros (infit=1.58, outfit = 1.70), con lo cual ambos fueron descartados. Por otro lado, también se hicieron análisis para determinar la correlación punto medida, donde ninguno presentó una puntuación baja. Asimismo, se determinó que ningún ítem afectó la unidimensionalidad del constructo. Finalmente, siguiendo con los criterios de Linacre (2022), se llegó a la conclusión de que no existían ítems que presentaran funcionamiento diferencial. Tomando en cuenta todo lo anterior, la escala de Intolerante quedó en definitiva con diez ítems en total.

Resultados finales de la escala

Tabla 75

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Intolerante

	Ítems	Personas
Fiabilidad	.99	.70
Separación	11.46	1.53

Fuente: elaboración propia

En la tabla anterior podemos observar que, con algunos ajustes, esta escala podría fortalecer los estadísticos de fiabilidad y separación para que estos muestren de forma más determinante que la escala es sensible para diferenciar a personas con diferentes niveles de intolerante y que podríamos esperar consistencia de estas inferencias, de igual forma se comprueba que existe una jerarquía en los ítems de la escala. A su vez, el Alfa de Cronbach obtenido fue de .70, puntaje que se puede interpretar como relativamente alto (Taber, 2017), y que confirma lo explicado anteriormente, sobre como la escala muestra potencial para detectar rasgos de intolerancia, pero se puede beneficiar de algunas modificaciones para fortalecer su confiabilidad.

Tabla 76

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Intolerante.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2091	48	-2.90	.92	.94	No
2	1623	38	-1.44	.94	.86	-1.95
3	507	12	-.40	1.00	1.04	.20
4	99	2	.06	1.52	1.83	1.74

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala de Intolerante se ajusta a todos los criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 77

Estadísticos de los ítems para la escala de Intolerante

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlaciones punto-medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM136	1.34	.11	1.32	1.23	3.48	1.76	.32	.18
ITEM156	-1.88	.07	.79	.80	-3.45	-3.40	.56	-.10
ITEM112	-1.15	.07	1.07	1.07	1.11	1.03	.51	-.11
ITEM115	-.54	.08	.80	.80	-3.11	-3.07	.53	.02
ITEM139	1.37	.11	1.32	1.08	3.38	.67	.41	.00
ITEM102	.90	.10	1.36	1.37	4.25	3.24	.38	.03
ITEM124	.37	.09	.95	.87	-.76	-1.63	.57	.06
ITEM160	.09	.08	.90	.83	-1.41	-2.39	.64	-.02
ITEM162	.17	.08	.82	.80	-2.75	-2.73	.58	.08
ITEM164	-.67	.08	1.25	1.26	3.52	3.52	.51	.34

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Intolerante indican que, en el caso de los ítems, el promedio del Infit fue de 1.06 (DT= .22). Para el ajuste global externo, el promedio del Outfit fue de 1.01 (DT=.21). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.00 (DT= .60) y el valor del Outfit de 1.01 (DT=.69). Con esto se puede concluir que ninguno de los ítems presenta índices que desajusten de los valores esperados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente. Por otro lado, solamente 9.95% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no se encuentran diferencias significativas entre las predicciones del modelo y los datos empíricos que se encontraron posterior a los análisis. Después de realizar las valoraciones de los datos de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.32 a 0.64) por lo tanto, se confirma que cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 42.9%. Asimismo, el primer componente tenía un valor propio de 1.58 lo que corresponde a un 9% de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

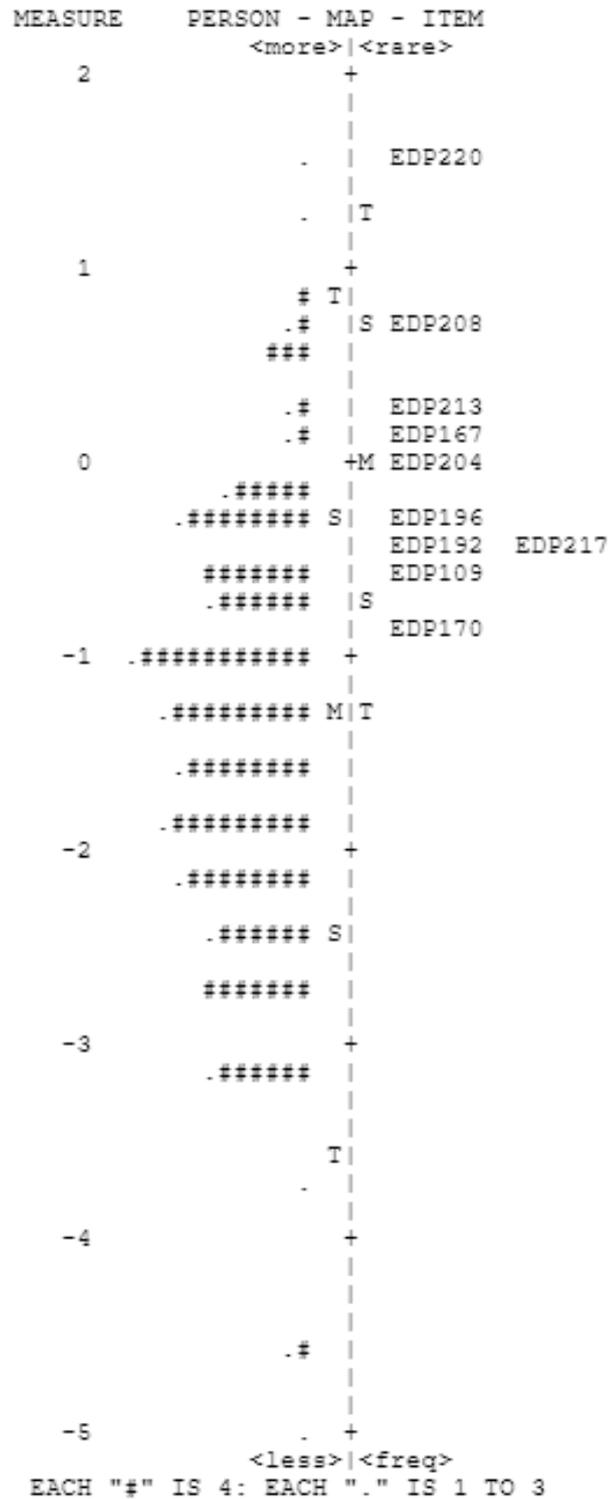
Los diez ítems que se mantuvieron en la escala han mostrado un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios predeterminados para detectar dicho criterio, y por ende no presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Intolerante que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y mostraron correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.5.4. Escala de Inflexible

Figura 27

Gráfico Escala de Inflexible – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Inflexible originalmente constaba de 11 ítems, después de realizar los análisis se decide eliminar un ítem del grupo. Siguiendo con los criterios de Linacre (2022), se detectó que el ítem 189 podría estar presentando un funcionamiento diferencial (DIF contrast = 0.52, $p < .001$), por lo que se reafirmó la decisión de eliminar este ítem.

También se realizaron análisis para determinar si existían ítems que desajustaron severamente el infit y/o el outfit, pero no fue el caso. De igual forma, se revisó la posibilidad de que los ítems presentaran una correlación punto medida baja, pero los datos no reflejaron que algún ítem presentara este problema. Finalmente, ningún ítem parece haber afectado la unidimensionalidad del constructo, con lo cual, la escala de Inflexible quedó en definitiva con diez ítems en total.

Resultados finales de la escala

Tabla 78

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Inflexible

	Ítems	Personas
Fiabilidad	.99	.77
Separación	8.30	1.83

Fuente: elaboración propia

Tal y como se observa en la Tabla 78, los estadísticos de fiabilidad y separación indican que la escala muestra criterios muy aceptables para indicar que la misma es sensible para diferenciar a personas con diferentes niveles de inflexibilidad y que podríamos esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. El Alfa de Cronbach de esta escala fue de .76, mismo que puede ser entendido como bastante alto (Taber, 2017), y que refuerza la explicación antes brindada.

Tabla 79

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Inflexible.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1381	32	-2.39	.95	.96	No
2	2045	47	-1.14	.93	.83	-2.14
3	762	18	-.22	.93	.92	.26
4	132	3	.00	1.42	1.65	1.87

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala de Inflexible alcanza todos los criterios predeterminados por Linacre (2002), mismos que corroboran la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 80

Estadísticos de los ítems para la escala de Inflexible

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM167	.21	.08	1.29	1.35	3.98	4.63	.48	.11
ITEM109	-.61	.07	1.05	1.05	.70	.82	.60	-.04
ITEM217	-.37	.08	1.13	1.12	1.90	1.79	.58	.00
ITEM220	1.51	.10	.79	.67	-3.12	-3.93	.64	.37
ITEM170	-.88	.07	1.12	1.16	1.86	2.33	.46	-.13
ITEM204	-.06	.08	1.08	1.11	1.19	1.60	.46	-.12
ITEM208	.67	.08	.93	.88	-.99	-1.73	.64	.00
ITEM213	.26	.08	.69	.71	-5.08	-4.76	.62	.11
ITEM192	-.42	.08	.78	.79	-3.61	-3.38	.64	-.15
ITEM196	-.30	.08	1.14	1.20	2.08	2.84	.48	.11

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Inflexible indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT= .18), mientras que para el ajuste global externo, el promedio del Outfit fue de 1.00 (DT=.22). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.03 (DT= .65) y el valor del Outfit de 1.00 (DT= .62). De esta forma, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente. Por otro lado, solamente 8.1% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos demuestran que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.46 a 0.64) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 37.4%. Asimismo, el primer componente tenía un valor propio de 1.56 lo que corresponde a un 9.8 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

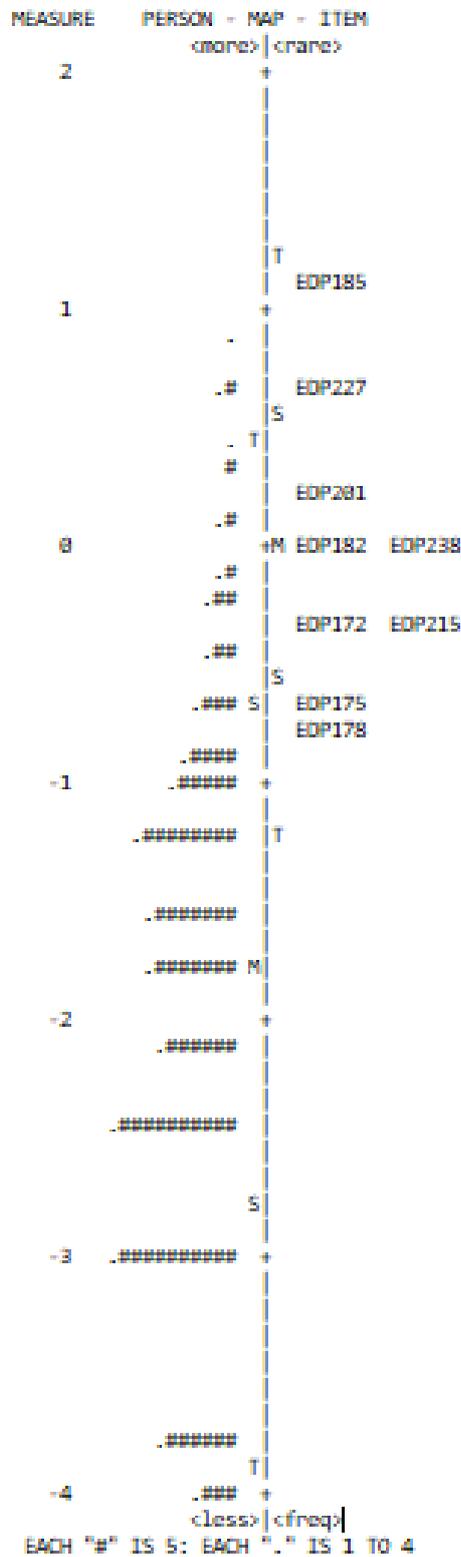
Los diez ítems que terminan conformando la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios esperados para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Inflexible que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.5.5. Escala de Falta de Planeamiento

Figura 28

Gráfico Escala Falta de Planeamiento – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Falta de Planeamiento originalmente contaba con 11 ítems, durante el proceso de ajuste y revisión se tomó la decisión de eliminar dos ítems. En una primera instancia, se eliminó el ítem 284 ya que desajustaba severamente tanto en el infit, como en el outfit (infit=1.96, outfit = 2.39). Posteriormente, al aplicar los criterios de Linacre (2022), se detectó otro ítem que debía ser eliminado, esto debido a que podía estar presentando funcionamiento diferencial, por lo que se procedió a eliminar el ítem 231 (DIF contrast = 0.52, $p < .001$) siendo de menor severidad para las mujeres de la muestra.

Por otro lado, se realizaron análisis para determinar si habían ítems que presentaron una correlación punto medida baja, pero este no fue el caso para esta escala. También se determinó que ningún ítem afectaba la unidimensionalidad del constructo. Finalmente, la escala de Falta de Planeamiento quedó en definitiva con nueve ítems en total.

Resultados finales de la escala

Tabla 81

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Falta de Planeamiento.

	Ítems	Personas
Fiabilidad	.98	.69
Separación	7.19	1.48

Fuente: elaboración propia

Tal y como se observa en la Tabla 81, los estadísticos de fiabilidad podrían ser aceptables con algunas mejores, mientras que la separación también puede beneficiarse de algunos ajustes en la escala para que la misma sea sensible para diferenciar a personas con diversos niveles de falta de planeamiento, y que de esta forma, se pueda esperar consistencia de estas inferencias, así como una comprobación de la existencia de una jerarquía en los ítems de la escala. En esta escala el Alfa de Cronbach obtenido fue de .76, valor que puede ser entendido como bastante alto según Taber (2017), y que confirma que la escala muestra niveles de confiabilidades aceptables y con potencial para una buena discriminación de los rasgos de Falta de Planeamiento, pero que podría verse fortalecida con algunos ajustes.

Tabla 82

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Falta de Planeamiento.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2035	52	-2.41	1.02	1.01	No
2	1293	33	-1.32	.93	.86	-1.49
3	461	12	-.48	.98	.95	.12
4	99	3	-.07	1.21	1.28	1.38

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala de Falta de Planeamiento cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 83

Estadísticos de los ítems para la escala de Falta de Planeamiento.

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM215	-.31	.07	.94	.98	-.88	-.25	.57	.09
ITEM172	-.36	.07	1.21	1.16	2.79	2.04	.58	.00
ITEM175	-.68	.07	1.17	1.22	2.39	2.98	.54	.12
ITEM178	-.75	.07	1.10	1.12	1.45	1.69	.54	-.08
ITEM182	-.02	.08	.72	.74	-4.19	-3.44	.62	-.19
ITEM201	.21	.08	.77	.73	-3.34	-3.35	.61	-.28
ITEM227	.71	.09	.82	.78	-2.29	-2.19	.53	-.12
ITEM238	.04	.08	1.30	1.16	3.69	1.86	.57	.48
ITEM185	1.16	.10	1.01	.88	.13	-.93	.49	-.10

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Falta de Planeamiento indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT= .19). Para el ajuste global externo, el promedio del Outfit fue de .97 (DT=.19). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.02 (DT= .55) y el valor del Outfit de .97 (DT= .54). De esta forma, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango recomendado para una medición adecuada del constructo. Por otro lado, solamente 6.71% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los resultados dan a entender que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Una vez realizado el análisis de los valores de correlación punto-medida se determinó que todos los ítems obtuvieron valores mayores a .3 (0.49 a 0.62) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para determinar la calidad del ajuste que presenta la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 37.5%. Asimismo, el primer componente obtuvo un valor propio de 1.87 lo que corresponde a un 13 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad esperados.

Los nueve ítems que se mantuvieron en la escala muestran un alto funcionamiento diferencial. Ninguno tiene un DIF por encima del rango esperado para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Falta de Planeamiento que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

Resultados iniciales de la escala

La escala de Inconstante originalmente constaba de 10 ítems, durante el proceso de ajuste y revisión se tomó la decisión de conservar la totalidad de los mismos. Esto debido a que los ítems no desajustaron severamente el infit y/o el outfit. Por otro lado, los ítems no presentaron una correlación punto medida baja, con lo cual no se hicieron descartes con base a este criterio. La unidimensionalidad tampoco se mostró afectada negativamente por los ítems de esta escala. Finalmente, según los criterios de Linacre (2022), no se detectaron ítems que podrían estar presentando funcionamiento diferencial. Tomando en cuenta lo descrito anteriormente, se toma la decisión de sostener la escala de Inconstante con sus diez ítems en total.

Resultados finales de la escala

Tabla 84

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Inconstante

	Ítems	Personas
Fiabilidad	.99	.78
Separación	10.90	1.87

Fuente: elaboración propia

Tal y como se observa en la Tabla 84, los estadísticos de fiabilidad y separación indican que la escala muestra potencial para ser sensible al diferenciar a personas con diferentes niveles de rasgos de Inconstante, y que podríamos esperar estabilidad de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. Adicionalmente, el valor del Alfa de Cronbach obtenido fue .79, mismo que se podría considerar bastante alto (Taber, 2017), y que reafirma la capacidad de la escala para diferenciar los diferentes niveles de Inconstante, y a su vez, muestran la posibilidad de mejorar la confiabilidad con algunos cambios.

Tabla 85

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Inconstante

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1757	41	-2.47	1.00	.99	No
2	1673	39	-1.13	.94	.84	-1.77
3	702	16	-.14	.95	.93	.20
4	188	4	.43	1.26	1.39	1.57

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala de Inconstante cumple con los criterios establecidos por Linacre (2002), mismos que garantizan la eficacia de las categorías de respuesta utilizadas en esta escala.

Tabla 86

Estadísticos de los ítems para la escala de Inconstante

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlaciones punto-medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM118	-1.00	.07	.83	.84	-2.73	-2.62	.63	-.12
ITEM119	1.64	.10	1.18	.83	1.98	-1.40	.52	-.20
ITEM97	-.60	.07	1.07	1.05	1.04	.76	.64	.00
ITEM99	-1.12	.07	1.31	1.30	4.47	4.28	.53	-.23
ITEM101	.27	.08	1.05	1.05	.76	.71	.49	.10
ITEM105	.22	.08	1.05	1.09	.79	1.15	.55	-.06
ITEM123	1.16	.09	1.09	.98	1.19	-.17	.51	.12
ITEM159	-.86	.07	.97	1.00	-.38	-.01	.62	.00
ITEM108	-.04	.08	.70	.76	-4.85	-3.58	.61	.10
ITEM216	.33	.08	1.00	.96	.02	-.45	.58	.32

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Inconstante indican que, en el caso de los ítems, el promedio del Infit fue de 1.03 (DT= .16). Para el ajuste global externo, el promedio del Outfit fue de .99 (DT=.15). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de 1.02 (DT= .59) y el valor del Outfit de .99 (DT= .60). De esta forma, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se evidencia que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente. Por otro lado, solamente 9.02% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.49 a 0.64) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 44.1%. Asimismo, el primer componente tenía un valor propio de 1.69 lo que corresponde a un 9.5 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

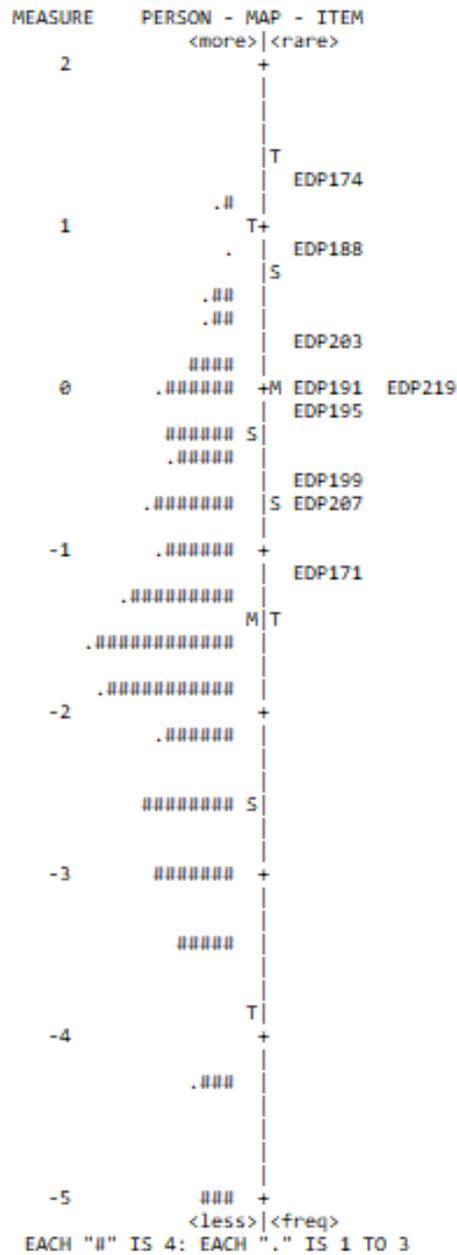
Los diez ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Inconstante que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.6.2. Escala de Poco Confiable

Figura 30

Gráfico Escala de Poco Confiable – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Poco Confiable, originalmente constaba de 10 ítems, durante el proceso de ajuste se tomó la decisión de eliminar uno de estos ítems. Inicialmente, se hicieron análisis para comprender si los ítems desajustaban severamente el infit y/o el outfit, pero este no fue

el caso. También, se revisó que los ítems no presentaran una correlación punto medida baja, rubro que también arrojó resultados positivos. A su vez, la unidimensionalidad no se vio afectada por este grupo de ítems. Finalmente, siguiendo con los criterios de Linacre (2022), se detectó que uno de los ítems podría estar presentando funcionamiento diferencial, por lo que se tomó la decisión de eliminar el ítem 224 (DIF contrast = 0.70, $p < ,001$) siendo de menor severidad para los hombres de la muestra. La escala de Poco Confiable quedó en definitiva con nueve ítems en total.

Resultados finales de la escala

Tabla 87

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Poco Confiable

	Ítems	Personas
Fiabilidad	.99	.77
Separación	8.70	1.85

Fuente: elaboración propia

Tal y como se observa en la Tabla 88, los estadísticos de fiabilidad y separación indican que la escala tiene potencial para ser sensible para diferenciar a personas con diferentes niveles de poca confiabilidad y que, con algunos ajustes, podríamos esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. Además, para esta escala se obtuvo un valor de Alfa de Cronbach de .78, el cual se puede considerar bastante alto (Taber, 2017), valor que confirma lo anteriormente expuesto sobre esta escala, teniendo la misma datos que indican una confiabilidad aceptable o buena, y que tiene potencial para solidificarse en este aspecto con algunas modificaciones.

Tabla 88

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Poco Confiable

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1457	37	-2.47	.98	.99	No
2	1664	43	-1.21	.93	.93	-2.02
3	636	16	-.11	.92	.92	.26
4	131	3	.18	1.27	1.38	1.75

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta muestran que la escala de Poco Confiable, cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 89

Estadísticos de los ítems para la escala de Poco Confiable

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM219	.06	.08	.81	.82	-2.99	-2.60	.61	-.06
ITEM203	.30	.08	1.17	1.12	2.34	1.58	.57	.28
ITEM207	-.70	.07	.75	.76	-4.11	-3.85	.65	-.36
ITEM188	.80	.09	1.07	1.08	.96	.98	.49	.49
ITEM191	.06	.08	.90	.92	-1.50	-1.10	.58	-.29
ITEM195	-.16	.08	.85	.86	-2.26	-2.02	.64	.13
ITEM199	-.56	.07	1.07	1.04	1.02	.64	.66	.00
ITEM171	-1.10	.07	1.26	1.29	3.74	4.17	.55	-.14
ITEM174	1.30	.10	1.20	1.11	2.45	1.05	.47	.14

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Poco Confiable indican que, en el caso de los ítems, el promedio del Infit fue de 1.01 (DT= .18). Para el ajuste global externo, el promedio del Outfit fue de 1.00 (DT=.16). En cuanto a estos indicadores de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT= .59) y el valor del Outfit de 1.00 (DT= .61). De esta forma, se confirma que ninguno de los ítems muestra índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se corrobora que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para una medición adecuada. A su vez, solamente el 6.71% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite determinar que el ajuste de las personas al modelo es admisible). De esta forma, los datos señalan que no existen discrepancias significativas entre las predicciones del modelo y los datos empíricos alcanzados. Posterior al análisis de los valores de correlación punto-medida se confirmó que todos los ítems obtuvieron valores mayores a .3 (0.47 a 0.66) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para determinar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 41.7%. Asimismo, el primer componente tenía un valor propio de 1.35 lo que corresponde a un 8.8 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad determinados.

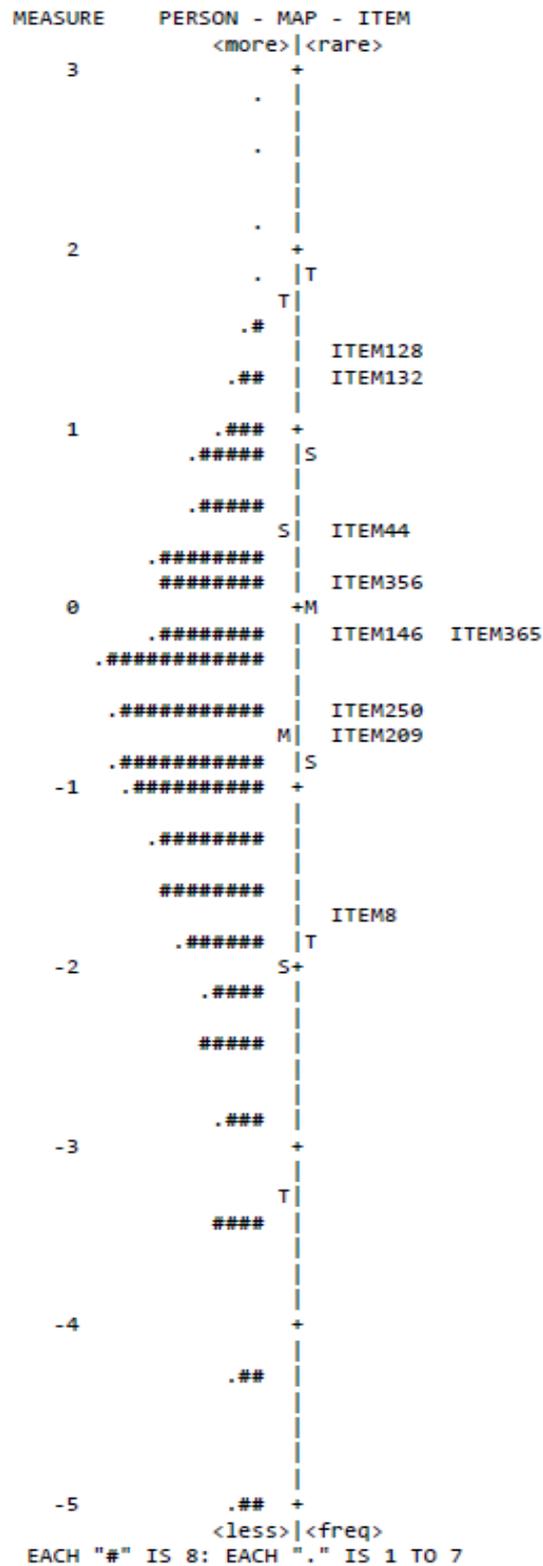
Los nueve ítems que se mantuvieron en la escala muestran un alto funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios predeterminados para detectar dicho criterio, ni se muestran sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Poco Confiable, mismos que mostraron un buen ajuste bajo el modelo de Rasch, no presentan DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.6.3. Escala de Imprudente

Figura 31

Gráfico Escala de Imprudente – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Imprudente originalmente constaba de 12 ítems, durante el proceso de ajuste se eliminaron los siguientes tres ítems. El ítem 290 desajustó severamente en el infit y el outfit, siendo los valores de infit=1.55, y el de outfit = 1.64. Posteriormente, siguiendo con los criterios de Linacre (2022), se detectaron dos ítems que mostraban señales de un funcionamiento diferencial, por lo que se toma la decisión de eliminar el ítem 287 (DIF contrast = 0.72, $p < ,001$) siendo de menor severidad para los hombres de la muestra y el ítem 237 (DIF contrast = 0.82, $p < ,001$) siendo de menor severidad para los hombres de la muestra.

Adicionalmente, se hicieron análisis para determinar si los ítems mostraban una correlación punto medida baja, pero no fue el caso. Finalmente, se confirmó que el grupo de ítems no afecta la unidimensionalidad. La escala de Imprudente quedó en definitiva con nueve ítems en total.

Resultados finales de la escala

Tabla 90

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Imprudente

	Ítems	Personas
Fiabilidad	.98	.73
Separación	6.60	1.66

Fuente: elaboración propia

Tal y como se observa en la Tabla 90, los estadísticos de fiabilidad y separación indican que la escala muestra potencial para ser sensible para diferenciar a personas con diferentes niveles de imprudencia y que con algunos ajustes, podríamos esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. Asimismo, el Alfa de Cronbach obtenido en esta escala fue de .73, un valor que puede ser considerado como alto (Taber, 2017), y que respalda que la escala muestra datos aceptables para discriminar este constructo, pero que puede fortalecerse con algunas mejoras.

Tabla 91

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Imprudente.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1361	35	-1.90	.99	1.02	No
2	1687	43	-.99	.92	.90	-1.67
3	666	17	-.21	.92	.90	.28
4	174	4	.16	1.24	1.32	1.40

Fuente: elaboración propia

Las estadísticas presentadas de las categorías de respuesta confirman que la escala Imprudente, cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 92

Estadísticos de los ítems para la escala de Imprudente

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM177	-.92	.07	1.22	1.32	3.23	4.61	.46	.33
ITEM181	-.22	.07	.93	.97	-1.05	-.47	.54	.14
ITEM200	.22	.07	.79	.77	-3.24	-3.40	.62	-.19
ITEM230	-.34	.07	.62	.63	-6.70	-6.34	.65	.05
ITEM234	.54	.08	.82	.78	-2.66	-3.10	.61	-.39
ITEM184	.46	.08	.84	.82	-2.44	-2.47	.56	-.12
ITEM283	.07	.07	.98	1.01	-.23	.10	.52	-.57
ITEM270	-.43	.07	1.49	1.48	6.49	6.34	.46	.31
ITEM274	.63	.08	1.33	1.31	4.21	3.61	.47	.28

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Imprudente indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT= .27). Para el ajuste global externo, el promedio del Outfit fue de 1.01 (DT=.28). En cuanto a los indicadores de las personas, el valor del ajuste interno conseguido fue de 1.02 (DT= .62) y el valor del Outfit de 1.01 (DT= .64). De esta forma, ninguno de los ítems arroja índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se demuestra que los valores de ajuste de todos los ítems se encuentran en el rango óptimo para medir adecuadamente. Adicionalmente, únicamente el 9.49% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.46 a 0.65) por lo tanto, se confirma que cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 34.8%. De igual forma, el primer componente tenía un valor propio de 1.80 lo que corresponde a un 13.1 % de la varianza explicada, lo que implica que este cumple con los criterios de unidimensionalidad establecidos.

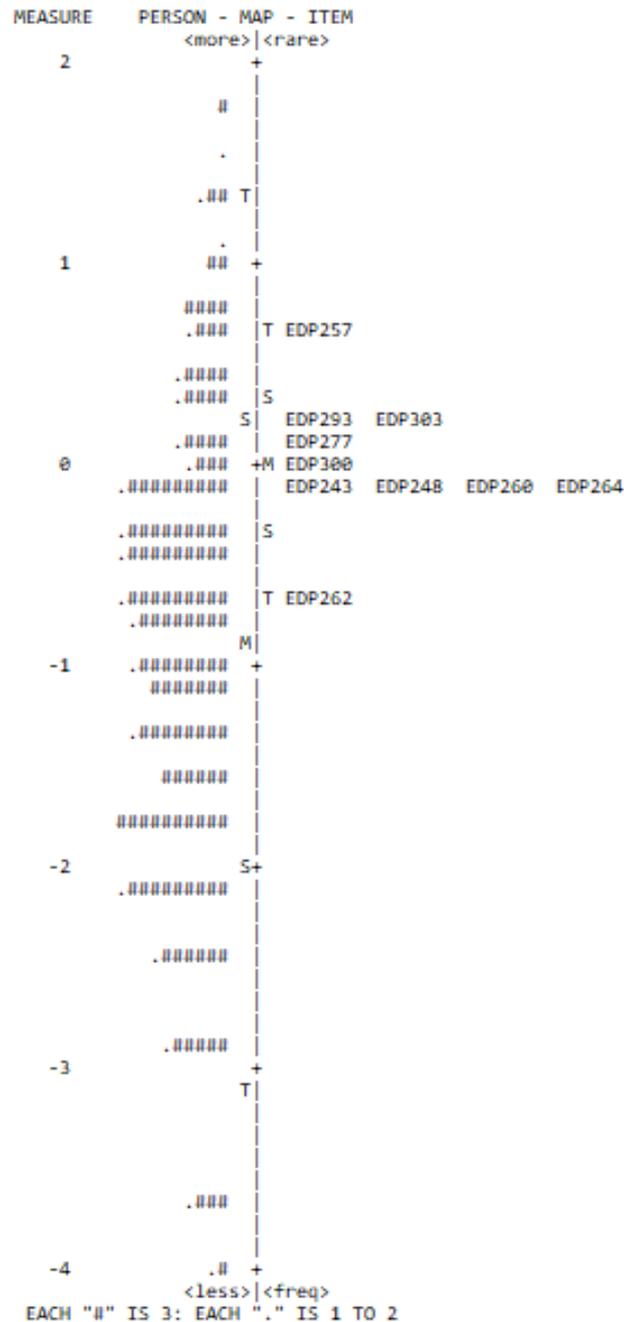
Los nueve ítems que se mantuvieron en la escala demuestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Imprudente que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.6.4. Escala de Intranquilo

Figura 32

Gráfico Escala de Intranquilo – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Intranquilo originalmente constaba de 10 ítems, durante el proceso de ajuste y revisión, se tomó la decisión de no eliminar ningún ítem. Esto debido a que ninguno

de ellos desajustó severamente el infit y/o el outfit. Por otro lado, los ítems tampoco presentaron una correlación punto medida baja. A su vez, la unidimensionalidad no se vio afectada por los ítems de esta escala. Finalmente, al correr análisis siguiendo los criterios de Linacre (2022), no se detectaron ítems que presentaran un funcionamiento diferencial. Con lo cual, se toma la decisión de mantener la escala de Intranquilo con diez ítems en total.

Resultados finales de la escala

Tabla 93

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala Intranquilo

	Ítems	Personas
Fiabilidad	.95	.81
Separación	4.59	2.06

Fuente: elaboración propia

Tal y como se observa en la Tabla 93, los estadísticos de fiabilidad y separación indican que la escala es sensible para diferenciar a personas con diferentes niveles de rasgos de intranquilidad y que podríamos esperar consistencia de estas inferencias, así como que se comprueba que existe una jerarquía en los ítems de la escala. Adicionalmente, el Alfa de Cronbach obtenido fue de .83, valor que puede ser considerado como robusto según Taber (2017), y que confirma la confiabilidad de la escala para detectar este constructo.

Tabla 94

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Intranquilo.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1456	34	-1.75	1.01	1.02	No
2	1565	36	-.81	.91	.90	-1.37
3	984	23	-.06	.94	.92	.00
4	315	7	.44	1.14	1.18	1.36

Fuente: elaboración propia

Los estadísticos mostrados anteriormente que corresponden a las categorías de respuesta, muestran que la escala de Intranquilo cumple con todos los criterios propuestos por Linacre (2002), mismos que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 95

Estadísticos de los ítems para la escala de Intranquilo

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacione s punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM277	.09	.07	.81	.80	-3.18	-3.04	.71	.00
ITEM293	.24	.07	.67	.65	-5.68	-5.34	.73	-.10
ITEM257	.63	.07	.95	.88	-.74	-1.54	.60	.28
ITEM260	-.16	.07	.79	.81	-3.57	-2.99	.65	.00
ITEM262	-.68	.07	1.11	1.12	1.68	1.85	.53	-.18
ITEM264	-.11	.07	1.12	1.14	1.84	2.00	.61	.10
ITEM243	-.09	.07	1.18	1.16	2.73	2.24	.55	.14
ITEM248	-.12	.07	1.26	1.24	3.71	3.27	.60	-.26
ITEM300	.00	.07	1.20	1.25	2.90	3.33	.51	.00
ITEM303	.20	.07	.94	.94	-.90	-.81	.60	.07

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Intranquilo indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT= .19). Para el ajuste global externo, el promedio del Outfit fue de 1.00 (DT=.20). Al analizar estos estadísticos, pero esta vez en las personas, se puede observar que el valor del ajuste interno obtenido fue de 1.01 (DT= .54) y el valor del Outfit de 1.00 (DT= .53). De esta forma, concluimos que ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se demuestra que los valores de ajuste de todos los ítems se incluyen en el rango óptimo para medir adecuadamente. Por otro lado, solamente 6.48% de las personas presentaron valores por encima a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las

personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.51 a 0.73) por lo tanto, se confirma que cumplen con el criterio de validez de contenido requerido.

Para determinar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 40%. Asimismo, el primer componente tenía un valor propio de 1.72 lo que corresponde a un 10.3 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

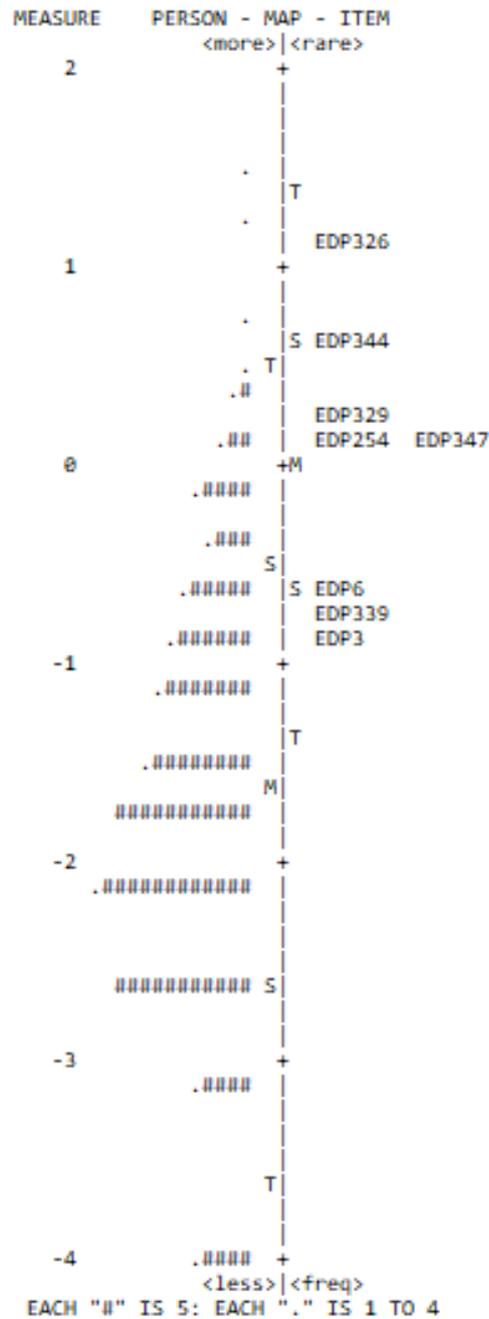
Los diez ítems que se mantuvieron en la escala muestran un alto funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Intranquilo, que mostraron tener un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.6.5. Escala de Disruptivo

Figura 33

Gráfico Escala Disruptivo – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Disruptivo originalmente constaba de 12 ítems, durante el proceso de ajuste se tomó la decisión de eliminar cuatro ítems. Inicialmente se realizaron análisis para

determinar cuáles ítems desajustaron severamente el infit y/o el outfit, los resultados arrojados mostraron que ninguno tenía un puntaje por encima de lo establecido, por lo cual no se descartó ninguno por este criterio. Por otro lado, los ítems tampoco presentaron una correlación punto medida baja. De igual forma, se logró establecer que ninguno de los ítems afectaba la unidimensionalidad de la escala. Finalmente, siguiendo con los criterios de Linacre (2022), se detectaron cinco ítems que podrían estar presentando funcionamiento diferencial, por lo que se toma la decisión de eliminar el ítem 307 (DIF contrast = 1.11, $p < ,001$) siendo de menor severidad para los hombres de la muestra y el ítem 343 (DIF contrast = 1.27, $p < ,001$) siendo de menor severidad para los hombres de la muestra. Finalmente se eliminó el ítem 304 (DIF contrast = 0.62, $p < ,001$), y el ítem 13 (DIF contrast = 0.84, $p < ,001$), ambos siendo de mayor severidad para los hombres de la muestra. La escala de Disruptivo quedó en definitiva con ocho ítems en total.

Resultados finales de la escala

Tabla 96

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Disruptivo

	Ítems	Personas
Fiabilidad	.99	.67
Separación	8.28	1.43

Fuente: elaboración propia

Tal y como se observa en la Tabla 96, los estadísticos de fiabilidad y separación indican que es necesario realizar cambios en la escala para que esta sea sensible para diferenciar a personas con diferentes niveles de rasgos disruptivos y que se pueda esperar consistencia de estas inferencias, de igual manera se comprueba que existe una jerarquía en los ítems de la escala. El Alfa de Cronbach obtenido en esta escala fue de .68, mismo que puede ser considerado razonable según Taber (2017), dato que respalda que la escala muestra potencial para detectar los rasgos del constructo de Disruptivo, pero que requiere algunas modificaciones para llegar a consolidarse.

Tabla 97

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Disruptivo.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	1490	43	-2.38	.96	.97	No
2	1436	42	-1.26	.97	.89	-1.82
3	430	12	-.28	.88	.82	.36
4	100	3	-.16	1.47	1.73	1.46

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta confirman que la escala de Disruptivo cumple con los cuatro criterios propuestos por Linacre (2002) que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 98

Estadísticos de los ítems para la escala de Disruptivo

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM254	.14	.08	1.16	1.13	2.12	1.67	.55	.18
ITEM3	-.90	.07	.75	.76	-4.22	-3.90	.57	-.45
ITEM6	-.65	.07	.79	.82	-3.42	-2.85	.50	.00
ITEM339	-.77	.07	1.35	1.35	4.83	4.78	.55	.20
ITEM344	.63	.09	1.07	.98	.91	-.19	.56	.21
ITEM347	.13	.08	.98	.95	-.20	-.64	.62	.16
ITEM326	1.14	.10	1.00	.90	.01	-1.01	.51	.08
ITEM329	.28	.08	1.11	1.08	1.57	-1.11	.47	-.29

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Disruptivo indican que, en el caso de los ítems, el promedio del Infit fue de 1.03 (DT= .18). Para el ajuste global externo, el promedio del Outfit fue de 1.00 (DT=.18). En cuanto a los datos estadísticos de las personas, el valor del ajuste interno obtenido fue de 1.01 (DT= .64) y el valor del Outfit de 1.00 (DT=.65). De esta forma, ninguno de los ítems presenta índices que desajusten (Infit u Outfit > 1.5), por lo que se confirma que los valores de ajuste de todos los ítems se incluyen en el rango deseado para medir adecuadamente. Por otro lado, solamente 7.40% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay diferencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Después de realizare al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.47 a 0.62) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

Para investigar el ajuste de la prueba al supuesto de unidimensionalidad, se realizó un Análisis de componentes principales de los residuos (ACPR). El porcentaje de la varianza explicada por la medida Rasch fue del 36.9%. Asimismo, el primer componente tenía un valor propio de 1.56 lo que corresponde a un 12.4 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

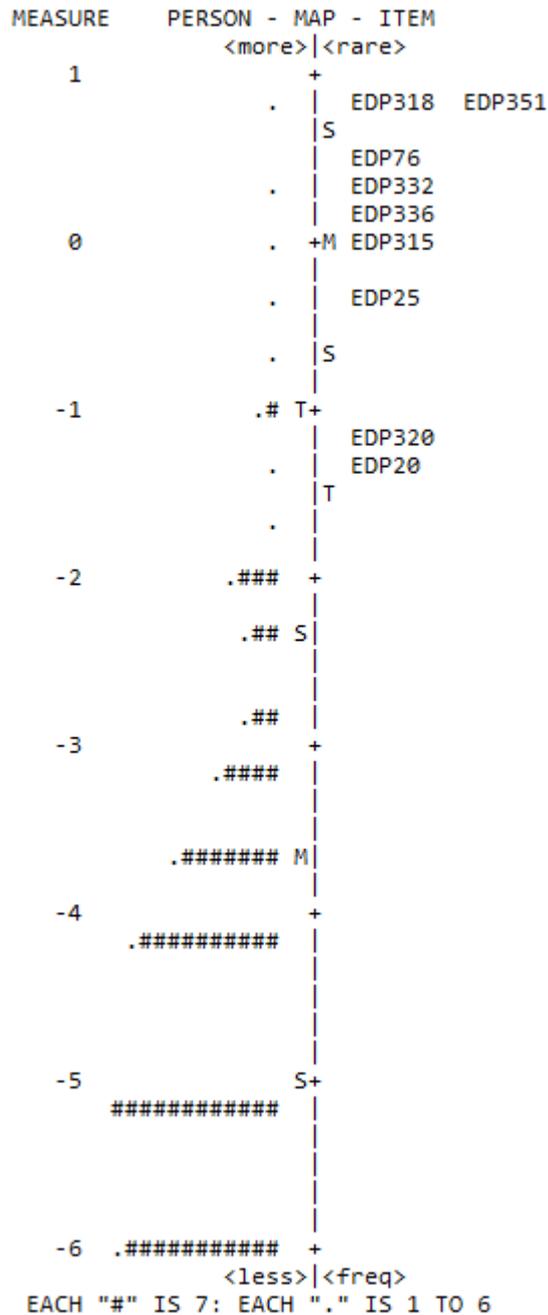
Los ocho ítems que se mantuvieron en la escala muestran un buen funcionamiento diferencial. Ninguno tiene un DIF por encima de los criterios establecidos para detectar dicho criterio, ni presenta sesgos por sexo.

La validez de constructo fue respaldada por todos los ítems en la escala de Disruptivo que tenían un buen ajuste bajo el modelo de Rasch, estaban libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

1.6.6. Escala de Agresivo

Figura 34

Gráfico Escala de Agresivo – Distribución casos e ítems



Resultados iniciales de la escala

La escala de Agresivo originalmente constaba de 9 ítems, durante el proceso de ajuste y revisión, se tomó la decisión de no eliminar ningún ítem. Sin embargo, siguiendo los

principios y recomendaciones de Linacre (2010), se decidió eliminar un total de 26 personas de la muestra, mismas que estaban desajustando severamente en la escala, y por ende afectando las estadísticas de la misma. Esta decisión logró un aumento de fiabilidad de .13 en las personas, y un aumento en la separación de .29 en las personas, y de .96 en los ítems.

No se encontraron desajustes severos en el infit ni en el outfit. De igual manera, los ítems no presentaron una correlación punto medida baja. El grupo de ítems tampoco afectó la unidimensionalidad del constructo. Finalmente, siguiendo con los criterios de Linacre (2022), se detectaron un ítem que podría estar presentando funcionamiento diferencial, el ítem 318 (DIF contrast = 0.60, $p < .001$) siendo de menor severidad para los hombres de la muestra; sin embargo, se optó por mantener el mismo debido a su contenido, ya que consideramos que es necesario que el mismo se aplicado a una muestra más amplia y heterogénea antes de tomar una decisión definitiva sobre si debe ser eliminado, o no. La escala de Agresivo terminó compuesta en definitiva por nueve ítems en total.

Resultados finales de la escala

Tabla 99

Estadísticos de Fiabilidad y Separación de ítems y personas en la escala de Agresivo

	Ítems	Personas
Fiabilidad	.97	.61
Separación	5.78	1.25

Fuente: elaboración propia

Tal y como se observa en la Tabla 99, los estadísticos de fiabilidad y separación indican que la escala puede beneficiarse de ajustes para obtener una mayor sensibilidad para diferenciar a personas con diferentes niveles de agresividad y que se pueda esperar consistencia de estas inferencias, además esto ayudaría a demarcar con mayor claridad la existencia una jerarquía en los ítems de la escala. Adicionalmente, el Alfa de Cronbach de esta escala fue de .80, mismo que puede ser considerado como bastante alto (Taber, 2017), y que reafirma el potencial de esta escala para dar resultados confiables al detectar rasgos de agresividad.

Tabla 100

Estadísticos correspondientes a la Estructura de Categorías (k=4) para la escala de Agresivo.

Categoría	Muestra	Frecuencia (%)	Promedio	InFit	OutFit	Paso
1	2611	71	-4.27	1.04	1.03	No
2	921	25	-2.66	.96	.91	-2.80
3	115	3	-1.06	1.07	1.02	.27
4	10	1	.08	1.15	1.19	2.54

Fuente: elaboración propia

Los estadísticos de las categorías de respuesta señalan que la escala de Agresivo, cumple con los criterios expuestos por Linacre (2002), mismos que garantizan la calidad de las categorías de respuesta utilizadas en esta escala.

Tabla 101

Estadísticos de los ítems para la escala de Agresivo

Ítem	Nivel de severidad (en logits)	Error estándar	MNSQ		ZTSD		Correlacion es punto- medida	DIF contrast
			Infit	Outfit	Infit	Outfit		
ITEM332	.41	.13	.92	.80	-.85	-1.65	.60	.23
ITEM351	.79	.14	1.08	1.01	.75	.13	.54	-.03
ITEM315	.00	.12	1.04	1.00	.54	.01	.61	-.11
ITEM318	.89	.15	1.01	.96	.13	-.19	.54	.26
ITEM336	.12	.13	.86	.84	-1.65	-1.51	.62	.08
ITEM20	-1.31	.10	1.10	1.10	1.29	1.33	.62	-.14
ITEM25	-.33	.12	1.15	1.11	1.76	1.22	.59	.07
ITEM320	-1.12	.11	1.01	1.01	.22	.14	.66	-.26
ITEM76	.55	.14	.88	.81	-1.26	-1.40	.59	-.07

Fuente: Elaboración propia

Los estadísticos de ajuste respectivos para la escala de Agresivo indican que, en el caso de los ítems, el promedio del Infit fue de 1.00 (DT= .09). Para el ajuste global externo, el promedio del Outfit fue de .96 (DT=.11). En cuanto a los indicadores de las personas, el valor del ajuste interno obtenido fue de .99 (DT= .39) y el valor del Outfit de .96 (DT= .54). De esta forma, ninguno de los ítems presenta índices de ajuste inadecuados (Infit u Outfit > 1.5), por lo que se confirma que los valores de ajuste de todos los ítems se incluyen en el rango deseado para medir adecuadamente. Por otro lado, después de eliminar parte de la muestra que desajustaba severamente, solamente el 4.92% de las personas presentaron valores mayores a 2 en Infit u Outfit (lo que implica un porcentaje adecuado que permite concluir que el ajuste de las personas al modelo es aceptable). De esta forma, los datos señalan que no hay divergencias significativas entre las predicciones del modelo y los datos empíricos obtenidos. Posterior al análisis de los valores de correlación punto-medida se identificó que todos los ítems obtuvieron valores mayores a .3 (0.54 a 0.66) por lo tanto, se evidencia cumplen con el criterio de validez de contenido requerido.

También se realizó un Análisis de componentes principales de los residuos (ACPR), para determinar el ajuste de la prueba al supuesto de unidimensionalidad. El porcentaje de la varianza explicada por la medida Rasch fue del 38.5%. Asimismo, el primer componente tenía un valor propio de 1.61 lo que corresponde a un 11.1 % de la varianza explicada, cumpliendo de esta forma con los criterios de unidimensionalidad establecidos.

Los nueve ítems que conforman la escala muestran un buen funcionamiento diferencial. Y solo uno de ellos muestra un DIF por encima de los criterios establecidos para detectar dicho criterio.

La validez de constructo fue respaldada por todos los ítems en la escala de Agresivo que tenían un buen ajuste bajo el modelo de Rasch, están casi en su totalidad libres de DIF y tenían correlaciones ítem-total que eran positivas y cercanas a sus valores esperados.

2. Evidencias de validez

2.1. Análisis factorial de las escalas

Se realizó un análisis factorial exploratorio (AFE) con el método de Máxima Verosimilitud y rotación Promax al total de las 33 escalas de la prueba construida, método es comúnmente utilizado y recomendado para análisis factoriales (Lloret et al., 2014).

La medida Kaiser-Meyer-Olkin de adecuación de muestreo (KMO) y la prueba de esfericidad de Bartlett indicaron valores que determinan la posibilidad de realizar el análisis factorial en las escalas. Los valores esperados para la prueba KMO son entre 0.5 y 1 y en la prueba de esfericidad de Bartlett, menores a 0.5 (Guerra, 2017), donde los valores superiores a 0.8 en la prueba KMO se consideran deseables (Lloret et al., 2014); de esta manera se concluye que los resultados obtenidos para estos estadísticos son óptimos, estos son expuestos en la Tabla 102.

Tabla 102

KMO y Test de esfericidad de Bartlett

Medida Kaiser-Meyer-Olkin de adecuación de muestreo		.965
Prueba de esfericidad de Bartlett	Aprox. Chi-cuadrado	10646
	gl	528
	Sig	< .001

Fuente: Elaboración propia

Ligado a estos datos, el AFE reveló 3 factores que, en conjunto, explican 56.4% de la varianza total (ver tabla 103).

Tabla 103

Varianza explicada por factores

Varianza explicada por factores

Factor	SC Cargas	% de la Varianza	% Acumulado
1	7.84	23.8	23.8
2	6.08	18.4	42.3
3	4.69	14.2	56.4

Fuente: Elaboración propia

La matriz y resultados obtenidos revelan que contrario a los 6 factores esperados, los ítems se agrupan en 3 factores, los resultados de los pesos factoriales para cada escala se

muestran en la Tabla 104. El primer factor, y el más grande en cuanto a cantidad de escalas, se compone de 14 de las 33 escalas, 5 de estas pertenecen al dominio Conductual, 2 al dominio Emocional, Sí Mismo, Cognitivo y Dominante, y 1 escala del dominio de Apego.

Por su parte, el segundo factor se compone de 9 escalas, 5 del dominio de Sí Mismo, 3 del dominio Dominante y 1 del dominio Cognitivo. Finalmente, se encuentra el tercer factor, compuesto de 9 escalas, 3 del dominio de Apego, 2 del Emocional y Cognitivo, y de los dominios Dominante y Conductual, 1 escala de cada uno.

Cabe destacar que las escalas de Autojustificación, Falta de Estabilidad Emocional, Sin Capacidad para Experimentar Placer e Indiferente, cargan tanto en el factor 1 como el factor 3 con pesos factoriales similares, pero más significativos en el factor 1; de igual manera, las escalas Sentido de Ser Único, Insincero y Manipulador, cargan al factor 1 y 2, siendo el mayor peso para las primeras dos en el factor 2, mientras que para Manipulador en el factor 1. Asimismo, la escala Falta de Ansiedad, del dominio Emocional, no carga en ningún factor.

Tabla 104*Matriz de estructura para las dimensiones*

Cargas de los Factores		Factor			Unicidad
		1	2	3	
Dominio					
5	Poco Confiable	0.893			0.227
5	Intranquilo	0.887			0.348
5	Inconstante	0.858			0.32
1	Autoconcepto Inestable	0.793			0.35
6	Falta de Concentración	0.771			0.375
5	Imprudente	0.744			0.419
6	Falta de Planeamiento	0.734			0.505
3	Engañoso	0.553			0.423
3	Manipulador	0.548	0.408		0.377
2	Falta de Estabilidad Emocional	0.489		0.358	0.469
4	Indiferente	0.487		0.354	0.312
1	Autojustificación	0.473		0.314	0.397
2	Sin Capacidad para Experimentar Placer	0.457		0.369	0.528
5	Disruptivo	0.38			0.696
1	Sentido de Ser Único	-0.318	1.059		0.3
1	Autoagrandecimiento		0.861		0.251
1	Centrado en Sí Mismo		0.839		0.283
3	Dominante		0.748		0.217
3	Charlatán		0.734		0.322
1	Sentido de Invulnerabilidad		0.628		0.569
3	Insincero	0.498	0.619		0.252
1	Sentido de Derecho		0.579		0.403
6	Inflexible		0.317		0.431
2	Falta de Ansiedad				0.765
3	Antagónico			0.735	0.365
4	Sin Empatía			0.712	0.563
4	Desapegado			0.615	0.579
6	Suspicious			0.606	0.635
4	No Adquiere Compromisos			0.54	0.471
2	Falta de Profundidad Emocional			0.517	0.673
6	Intolerante		0.348	0.447	0.427

2	Falta de Remordimiento	0.426	0.613
5	Agresivo	0.309	0.518

Método de extracción: Máxima Verosimilitud.

Método de rotación: Promax con normalización Kaiser.

*Dominios enumerados: Sí Mismo (1), Emocional (2), Dominancia (3), Apego (4), Conductual (5) y Cognitivo (6). Fuente: Elaboración propia

2.2. Correlaciones

A continuación, se presentan los resultados de las correlaciones de Pearson para evaluar la validez de criterio de los dominios de la prueba construida. Es importante notar que debido a la varianza de método compartida y como esta puede inflar los coeficientes de correlación a un grado desconocido, solamente se interpretaron los tamaños de efectos moderados o altos ($r = .40$ en adelante) como significativos (Sellbom, et al., 2019).

Tabla 105

Correlaciones de dominios con escalas del MMPI-2. N=59-60

	D. Sí Mismo	D. Emocional	D. Dominancia	D. Apego	D. Conductual	D. Cognitivo
Dp	0.313 p=.016	0.443 p<.001	0.369 p=.004	0.347 p=.007	0.577 p<.001	0.401 p=.002
Ma	0.386 p=.003	0.406 p=.001	0.451 p<.001	0.329 p=.010	0.521 p<.001	0.426 p<.001
AGR	0.412 p=.001	0.230 p=.077	0.398 p=.002	0.203 p=.119	0.212 p=.104	0.206 p=.118
DIS	0.465 p<.001	0.505 p<.001	0.545 p<.001	0.518 p<.001	0.577 p<.001	0.533 p<.001
RC4	0.450 p<.001	0.553 p<.001	0.549 p<.001	0.525 p<.001	0.709 p<.001	0.562 p<.001
RC9	0.525 p<.001	0.389 p=.002	0.631 p<.001	0.362 p=.005	0.632 p<.001	0.506 p<.001
PAS	0.458 p<.001	0.591 p<.001	0.585 p<.001	0.605 p<.001	0.636 p<.001	0.635 p<.001
Antisocial	0.512 p<.001	0.562 p<.001	0.621 p<.001	0.573 p<.001	0.686 p<.001	0.651 p<.001

Fuente: Elaboración propia

Los datos de la tabla 105 evidencian una fuerte correlación entre los diferentes dominios del instrumento con las escalas del MMPI-2. Con la escala Antisocial, los dominios de Sí Mismo, Emocional y Apego correlacionan de forma moderada, y los dominios de Dominancia, Conductual y Cognitivo lo hacen de forma alta. PAS presenta correlaciones moderadas (valores cercanos a altas) y altas con los dominios Emocional, Dominancia,

Apego, Conductual y Cognitivo. RC 9 mostró moderada correlación con Sí Mismo y Cognitivo, y alta con Dominancia y Conductual. Mientras que RC4, alto con Conductual, y moderadamente con Emocional, Dominancia, Apego y Cognitivo.

DIS correlaciona moderadamente con todos los dominios a excepción del Sí Mismo, presentando una correlación baja pero significativa. Por otra parte, Dp y Ma solamente correlacionan con el dominio Conductual de forma Moderada; y cabe resaltar que AGR no correlaciona significativamente con ningún dominio.

Tabla 106

Correlaciones de dominios con escalas del PAI. N=58

	D. Sí Mismo	D. Emocional	D. Dominancia	D. Apego	D. Conductual	D. Cognitivo
MAN-A	0.575 p<.001	0.484 p<.001	0.569 p<.001	0.412 p=0.001	0.667 p<.001	0.561 p<.001
MAN-G	0.359 p=.006	-0.118 p=.377	0.108 p=.420	-0.152 p=.255	-0.047 p=.727	-0.047 p=.729
ANT-A	0.453 p<.001	0.484 p<.001	0.509 p<.001	0.487 p<.001	0.553 p<.001	0.477 p<.001
ANT-E	0.573 p<.001	0.452 p<.001	0.510 p<.001	0.448 p<.001	0.516 p<.001	0.632 p<.001
ANT-B	0.428 p<.001	0.372 p=.004	0.435 p<.001	0.340 p=.009	0.601 p<.001	0.554 p<.001
AGR-A	-0.305 p=.020	-0.173 p=.194	-0.347 p=.008	-0.256 p=.052	-0.178 p=.182	-0.258 p=.052
AGR-V	0.118 p=.377	0.106 p=.430	0.284 p=.031	0.012 p=.928	0.214 p=.107	0.187 p=.163
AGR-F	0.488 p<.001	0.569 p<.001	0.569 p<.001	0.480 p<.001	0.700 p<.001	0.643 p<.001
DOM	0.181 p=.174	-0.203 p=.126	0.078 p=.563	-0.226 p=.088	-0.203 p=.127	-0.241 p=.071
WRM	0.096 p=.473	-0.214 p=.106	0.042 p=.755	-0.325 p=.013	0.118 p=.378	-0.009 p=.948
MAN	0.694 p<.001	0.376 p=.004	0.557 p<.001	0.276 p=.036	0.523 p<.001	0.464 p<.001
ANT	0.600 p<.001	0.544 p<.001	0.602 p<.001	0.531 p<.001	0.685 p<.001	0.686 p<.001
AGR	0.231 p=.081	0.352 p=.007	0.362 p=.005	0.189 p=.159	0.504 p<.001	0.399 p=.002
Antagonismo	0.478 p<.001	0.234 p=.077	0.400 p=.002	0.185 p=0.164	0.378 p=0.003	0.330 p=.012
Desinhibición	0.357 p=.006	0.553 p<.001	0.450 p<.001	0.429 p<.001	0.621 p<.001	0.598 p<.001

Fuente: Elaboración propia

Como se evidencia en la tabla 106, existe una correlación significativa del dominio Sí Mismo con las escalas (alta) y subescalas (moderada) MAN, ANT, MAN-A y ANT-E, del PAI. El dominio Emocional se correlaciona moderadamente solo con AGR-F, ANT y Desinhibición; por su parte el dominio de Apego lo hace con la escala ANT. El dominio Dominancia se correlaciona moderadamente con las subescalas MAN-A, ANT-A, ANT-E, AGR-F y la escala MAN, y de alto con ANT.

El dominio Conductual presenta correlaciones moderadas con ANT-A, ANT-E, MAN y AGR, y correlaciones altas con MAN-A, ANT-B, AGR-F, ANT y Desinhibición. Similarmente, el dominio Cognitivo se correlaciona con MAN-A, ANT-B y Desinhibición de forma moderada, y con ANT-E, AGR-F, y ANT, de forma alta.

Tabla 107

Correlaciones de dominios con escalas del UPPS-P. N=140-141

	Urgencia Negativa	Premeditación	Perseverancia	Búsqueda de Sensaciones	Urgencia Positiva	UPPS -P
D. Sí Mismo (N=140)	0.440 p<.001	0.121 p=.155	0.145 p=.087	0.457 p<.001	0.564 p<.001	0.527 p<.001
D. Emocional	0.412 p<.001	0.186 p=.027	0.232 p=.003	0.388 p<.001	0.493 p=.002	0.507 p<.001
D. Dominancia	0.493 p<.001	0.244 p=.004	0.241 p=.004	0.377 p<.001	0.550 p<.001	0.561 p<.001
D. Apego	0.393 p<.001	0.251 p=.003	0.272 p=.001	0.361 p<.001	0.518 p=.002	0.526 p=.002
D. Conductual	0.709 p<.001	0.425 p<.001	0.503 p<.001	0.436 p<.001	0.741 p<.001	0.806 p<.001
D. Cognitivo (N=140)	0.591 p<.001	0.272 p=.001	0.437 p<.001	0.335 p<.001	0.593 p<.001	0.634 p<.001

Fuente: Elaboración propia

En la tabla 107, se presentan las correlaciones con el UPPS-P, entre estas destacan las correlaciones altas del dominio Conductual con las escalas de Urgencia Negativa, Urgencia Positiva y todo el UPPS-P en conjunto, y una correlación moderada con Perseverancia. Similarmente, el dominio Cognitivo se correlaciona moderadamente con puntuaciones cercanas al rango alto, con Urgencia Negativa y Positiva, y de manera alta con el instrumento completo. Los dominios de Dominancia y Sí Mismo, correlacionan moderadamente con la escala de Urgencia Positiva y el UPPS-P total, mientras que el dominio emocional solo presenta esta última correlación.

Tabla 108*Correlaciones de dominios con escalas del IDPP. N=70-71*

	D. Sí Mismo	D. Apeg o	D. Emociona l	D. Dominanci a	D. Conductu al	D. Cognitiv o
B. Atención	0.379 p=.001	-0.039 p=.74	0.015 p=.89	0.431 p<.001	0.381 p=.001	0.214 p=.07
Manipulación	0.589 p<.001	0.500 p<.001	0.517 p<.001	0.717 p<.001	0.582 p<.001	0.535 p<.001
Engaño	0.494 p<.001	0.533 p<.001	0.571 p<.001	0.689 p<.001	0.654 p<.001	0.610 p<.001
Grandiosidad	0.595 p<.001	0.377 p=.001	0.431 p<.001	0.572 p<.001	0.412 p<.001	0.504 p<.001
Irresponsabilidad	0.273 p=.02	0.358 p=.002	0.353 p=.003	0.441 p<.001	0.667 p<.001	0.539 p<.001
Impulsividad	0.436 p<.001	0.339 p=.004	0.373 p=.001	0.493 p<.001	0.667 p<.001	0.530 p<.001
B. Sensaciones	0.456 p<.001	0.414 p<.001	0.415 p<.001	0.429 p<.001	0.560 p<.001	0.424 p<.001
Distractibilidad	0.199 p=.09	0.175 p=.14	0.162 p=.17	0.352 p=.003	0.561 p<.001	0.393 p<.001
P. Adicción	0.443 p<.001	0.423 p<.001	0.488 p<.001	0.459 p<.001	0.594 p<.001	0.511 p<.001
Insensibilidad	0.367 p=.002	0.596 p<.001	0.533 p<.001	0.472 p<.001	0.429 p<.001	0.541 p<.001
Desinhibición	0.466 p<.001	0.458 p<.001	0.466 p<.001	0.549 p<.001	0.763 p<.001	0.605 p<.001
Antagonismo	0.629 p<.001	0.504 p<.001	0.535 p<.001	0.755 p<.001	0.647 p<.001	0.619 p<.001
Desapego	0.314 p=.008	0.645 p<.001	0.603 p<.001	0.321 p=.006	0.244 p=.04	0.421 p<.001
Afecto Negativo	0.264 p=.02	0.253 p=.03	0.399 p=.004	0.360 p=.002	0.580 p<.001	0.453 p<.001
Psicoticismo	0.674 p<.001	0.644 p<.001	0.640 p<.001	0.669 p<.001	0.694 p<.001	0.702 p<.001

Fuente: Elaboración propia

Como se puede observar en la tabla 108, uno de los principales resultados obtenidos es la correlación alta y significativa existente entre la escala de Psicoticismo y los dominios del instrumento creado. De igual forma, existe una correlación alta, y además esperada, entre la escala de Desapego, y los dominios de Apego y Emocional. También se mostraron altas

correlaciones entre la escala de Antagonismo y los dominios Conductual, de Dominancia y Cognitivo.

Además, se correlacionó de manera significativa y alta la escala de Desinhibición con el dominio Conductual y el Cognitivo. Mientras que, en las escalas de Engaño y Manipulación, destaca su alta correlación con el Dominio de Dominancia. Finalmente, las escalas de Impulsividad e Irresponsabilidad, mostraron una alta correlación con el dominio conductual, y una correlación moderada con el cognitivo.

Algo destacable sobre estos resultados es que, si bien todas las escalas del IDPP mostraron al menos una correlación moderada con alguno de los dominios del instrumento creado, ninguno de los dominios mostró correlaciones altas con todas las escalas del IDPP, y a su vez, existieron numerosas correlaciones bajas. Esto es de interés y de valor, ya que es una clara señal de validez discriminante.

Tabla 109

Correlaciones de dominios con escalas del NEO PI-R. N=59-60

	D. Sí Mismo	D. Apeg o	D. Emociona l	D. Dominanci a	D. Conductua l	D. Cognitiv o
Neuroticismo	0.340 p=.008	0.370 p=.004	0.488 p<.001	0.457 p<.001	0.658 p<.001	0.556 p<.001
Extraversión	0.204 p=.01	-0.259 p=.04	-0.218 p=.09	0.109 p=.40	0.101 p=.44	-0.083 p=.53
Apertura	0.181 p=.16	-0.197 p=.13	-0.066 p=.61	0.073 p=.58	0.096 P=.46	-0.032 p=.80
Amabilidad	-0.562 p<.001	-0.587 p<.001	-0.575 p<.001	-0.598 p<.001	-0.501 p<.001	-0.467 p<.001
Responsabilidad	-0.285 p=.02	-0.445 p<.001	-0.489 p<.001	-0.440 p<.001	-0.650 p<.001	-0.520 p<.001

Fuente: Elaboración propia

Para culminar, son de gran interés las correlaciones del instrumento con el NEO PI-R, ya que este, a excepción del Factor de Neuroticismo, nos brinda evidencias para la validez discriminante. Como se puede observar en la tabla 109, existe una correlación fuerte entre el

Dominio Conductual y el Factor de Neuroticismo del NEO PI -R, mientras que el Dominio Cognitivo muestra una correlación moderada con este factor.

Adicionalmente, el Factor de Amabilidad mostró correlaciones negativas moderadas con el Dominio de Sí mismo, Dominio de Apego, y Dominio Emocional, mientras que, con el Dominio de Dominancia, si bien la correlación se encuentra dentro del rango que se podría considerar moderado, el puntaje está sumamente cercado a poder ser considerado como una correlación negativa alta.

Finalmente, el Factor de Responsabilidad reveló correlaciones negativas con dos de los dominios del instrumento. En el caso del Dominio Cognitivo se registró una correlación moderada, mientras que, con el Dominio Conductual, se evidenció una correlación alta.

2.3. Pruebas T Student

Se realizó una prueba t para muestras independientes con el fin de identificar si existían diferencias significativas en la población de acuerdo con su sexo (ver tabla 110), los resultados muestran que hubo diferencias en uno de los 6 dominios.

Específicamente el dominio de Apego, en promedio, los hombres reportaron mayores puntuaciones ($M = 83.5$, $SE = 11$) que las mujeres ($M=80.2$, $SE= 9.59$). Esta diferencia fue significativa [$t(430) = -3.345$, $p < .001$]. El tamaño del efecto fue medio, de $d= -0,32$. No se encontraron diferencias estadísticamente significativas para el resto de los dominios.

Tabla 110

Tabla comparativa de dominios: Medias, desviaciones estándar, prueba t y significancia bilateral para la muestra según género.

Dominio	Mujeres (n=250-253)		Hombres (n=178-179)		t	p	d
	Media	DE	Media	DE			
Sí Mismo	130.6	21.5	132.6	24.1	-9.21	.358	-0.0909
Emocional	95	11.8	94.9	12	0.169	.866	0.0165
Dominancia	99.7	18.47	103.4	20.9	-1.916	.056	-0.1891
Conductual	104	16.9	106.1	18.8	-1.186	.237	-0.1171
Cognitivo	99.6	13.79	101	15.1	-0.983	.326	-0.0962
Apego	80.2	9.59	83.5	11	-3.345	.001	-0.3267

Fuente: Elaboración propia

2.4. ANOVAS

Para los siguientes análisis, las edades se recodificaron para hacer tres grupos, con el fin de establecer ciertas comparaciones en las medias de los dominios. Los grupos etarios fueron 18 a 30 años, 31 a 45 años y de 46 años en adelante.

Tabla 111

Resultados ANOVA por grupo de edad según cada dominio.

DOMINIO		Suma de Cuadrados	gl	Media Cuadrática	F	p
COGNITIVO	Grupo de edad	1610	2	805	2.64	0.073
	Residuos	76211	250	305		
CONDUCTUAL	Grupo de edad	6181	2	3090	6.14	0.002
	Residuos	126277	251	503		
APEGO	Grupo de edad	1969	2	984	4.57	0.011
	Residuos	54024	251	215		
DOMINANCIA	Grupo de edad	7192	2	3596	7.05	0.001
	Residuos	127964	251	510		
EMOCIONAL	Grupo de edad	1011	2	506	1.77	0.172
	Residuos	71696	251	286		
SI MISMO	Grupo de edad	7573	2	3786	5.61	0.004
	Residuos	168800	250	675		

Fuente: elaboración propia

Como se puede observar en la Tabla 111 existieron diferencias significativas en cuatro de los seis dominios investigados en este estudio. Siendo estos el dominio Conductual ($p=.002$), de Apego ($p=.011$), de Dominancia ($p=.001$), y de Sí Mismo ($p=.004$). Mientras que, por su parte, los dominios Cognitivos y Emocionales, no mostraron efectos significativos.

Tabla 112*Comparaciones Post Hoc - Grupo de edad del dominio Conductual*

Comparación		Diferencia de Medias	EE	gl	t	P _{tukey}	d de Cohen
Grupo de edad	Grupo de edad						
DE 18 A 30 AÑOS	- DE 31 A 45 AÑOS	6.43	3.59	251	1.79	0.174	0.287
	- DE 46 Y MÁS AÑOS	12.7	3.79	251	3.35	0.003	0.566
DE 31 A 45 AÑOS	- DE 46 Y MÁS AÑOS	6.27	4.57	251	1.37	0.356	0.28

Nota. Las comparaciones se basan en medias marginales estimadas

Fuente: elaboración propia

Como se puede observar en la Tabla 112, existe una diferencia muy significativa ($p=.003$), entre el grupo de 18 a 30 años, y el de 46 y más años. Existiendo una diferencia de medias de 12.70, siendo la media del grupo menor de 110.7 y la del grupo mayor de 98. Adicionalmente, el tamaño del efecto en esta comparación se puede considerar como muy grande ($r=.566$), lo cual implica que este explica más del 25% de la varianza mostrada.

Tabla 113*Comparaciones Post Hoc - Grupo de edad del dominio de Apego*

Comparación		Diferencia de Medias	EE	gl	t	P _{tukey}	d de Cohen
Grupo de edad	Grupo de edad						
DE 18 A 30 AÑOS	- DE 31 A 45 AÑOS	-2.42	2.35	251	-1.03	0.557	-0.165
	- DE 46 Y MÁS AÑOS	6.25	2.48	251	2.52	0.033	0.426
DE 31 A 45 AÑOS	- DE 46 Y MÁS AÑOS	8.67	2.99	251	2.9	0.011	0.591

Nota. Las comparaciones se basan en medias marginales estimadas

Fuente: elaboración propia

Como se puede observar en la tabla anterior, existen diferencias significativas ($p=.033$) entre el grupo de 18 a 30 años, y el de 46 y más años, con una diferencia de medias del 6.25. Siendo la media del grupo menor de 70.1 y la del mayor de 63.8. En este caso, el

tamaño del efecto fue medio ($r=.426$), lo cual implica que este explica un poco más del 9% de la varianza.

Asimismo, existen diferencias significativas ($p=.011$), entre el grupo de 31 a 45 años, y el grupo de 65 años en adelante. Entre estos hubo una diferencia de medias del 8.67, ya que la media del grupo de 31 a 45 años fue de 72.5. Para esta comparación el tamaño del efecto fue grande ($p=.591$), explicando así más del 25% de la varianza observada.

Tabla 114

Comparaciones Post Hoc - Grupo de edad del dominio de Dominancia

Comparación		Diferencia de Medias	EE	gl	t	p _{tukey}	d de Cohen
Grupo de edad	Grupo de edad						
DE 18 A 30 AÑOS	- DE 31 A 45 AÑOS	6.99	3.61	251	1.94	0.131	0.31
	- DE 46 Y MÁS AÑOS	13.68	3.82	251	3.58	0.001	0.606
DE 31 A 45 AÑOS	- DE 46 Y MÁS AÑOS	6.69	4.6	251	1.46	0.314	0.296

Nota. Las comparaciones se basan en medias marginales estimadas

Fuente: elaboración propia

Tal y como se observa en la Tabla 114, existe una diferencia muy significativa, la más alta entre todos los ANOVAS realizados ($p=.001$), esto entre los grupos de 18 a 30 años, y los de 46 y más años. Entre ellos existe una diferencia de medias del 13.68, siendo la media del grupo menor 105.3, mientras que la del grupo de 46 en adelante del 91.6. A su vez, el tamaño el efecto que arrojan los datos es grande ($p=.606$), con lo cual se puede concluir que este explica más del 25% de la varianza.

Tabla 115*Comparaciones Post Hoc - Grupo de edad del dominio de Sí Mismo*

Comparación		Diferencia de Medias	EE	gl	t	P _{tukey}	d de Cohen	
DE 18 A 30 AÑOS	-	DE 31 A 45 AÑOS	5.43	4.16	250	1.3	0.394	0.209
	-	DE 46 Y MÁS AÑOS	14.53	4.4	250	3.31	0.003	0.559
DE 31 A 45 AÑOS	-	DE 46 Y MÁS AÑOS	9.1	5.29	250	1.72	0.199	0.35

Nota. Las comparaciones se basan en medias marginales estimadas

Fuente: elaboración propia

En la tabla 115, podemos encontrar una diferencia muy significativa ($p=.003$), entre el grupo de 18 a 30 años, y el grupo de 46 y más años, con una diferencia de medias de 14.53. Ya que la media del grupo más joven es de 143, mientras que el grupo de 46 el puntaje de media fue de 129. Asimismo, el tamaño del efecto fue grande ($p=.559$), con lo cual se puede afirmar que este explica más del 25% de la varianza observada.

Discusión de resultados

El constructo de la psicopatía ha sido objeto de controversia con el pasar de los años, desde cuestionamientos sobre su existencia, su presencia en ciertos grupos poblacionales, hasta la veracidad de su medición. Tal es el caso de Hart et al (2002), quienes en su momento cuestionaron con contundencia la existencia de la psicopatía en adultos jóvenes, tal y como se conoce en grupos adultos, y concluyeron que, en los test de su medición, solo se podía determinar que medían algo, pero no el qué, y que era imposible adjudicar con firmeza que el constructo de fondo era la psicopatía.

Adicionalmente, el constructo de la psicopatía ha tenido que recorrer un largo camino para liberarse de los estigmas de criminalización que ha cargado con el pasar de las décadas, mismo que es normal ya que su construcción y entendimiento emergió de la población penitenciaria. Sin embargo, el peso de ser una psicopatología asociada con criminales, ha

hecho más cuestionable su validez como constructo y su generalización a la población normativa.

Es por ello, que el modelo del CAPP propuesto por Cooke et al., (2012) representa un gran avance en el estudio de la psicopatía, ya que lleva este constructo a dimensiones generales, presentes en todas las personas y alejándose de supuestos más criminales como los que se podrían encontrar en el PCL-R donde se miden factores como “delincuencia juvenil” “revocación de libertad condicional” y “versatilidad criminal” (Hare, 2003). De esta manera el CAPP centra su atención en la medición específica de estas dimensiones puntuales, mismas que no son dependientes de escenarios como la historial criminal.

Con los resultados obtenidos de las escalas creadas en esta investigación, podemos afirmar que se han tomado pasos firmes y alentadores para el estudio de la psicopatía en la región, bajo un modelo que se aparta de lo categórico, y que ha mostrado solvencia en la dimensionalidad de sus constructos. Tal y como se mostró en el apartado de análisis de resultados, la mayoría de las escalas de esta prueba lograron alcanzar un puntaje mayor a 3 en separación de ítems, lo cual significa bajo el modelo de Rasch, que se logró comprobar la jerarquía en severidad de estos. Sumado a esto, todos los ítems de las escalas, a excepción de los ítems 308, 318, 333 y 348 que se mantuvo por contenido y necesidad de exploración con otras muestras, están libres de DIF, lo cual significa que los ítems de las escalas no tienen un sesgo que lleve a un subgrupo de la muestra a puntuar de forma significativa en comparación al otro, cuando ambos muestran el mismo nivel de severidad en el rasgo evaluado.

Adicionalmente, en el análisis de separación, pero en personas, la mayoría de las escalas estuvieron cerca, o superaron el puntaje recomendado de 2, lo cual implica que las escalas están mostrando un potencial, o han demostrado solventemente su sensibilidad para captar y distinguir entre desempeños altos y bajos de las personas con los ítems. Es probable que el haber tenido acceso solamente a población normativa repercutiera en una menor variabilidad en los niveles de severidad detectados con las distintas escalas. En la mayoría de las escalas el nivel de los ítems estaba por encima del nivel de las personas, lo que es esperable a la hora de medir constructos con altos grados de desajuste de la personalidad en población normativa.

Sumado a lo anterior, uno de los hallazgos más importantes en el análisis de datos, es que todas las escalas creadas para esta prueba, lograron cumplir los parámetros de

unidimensionalidad recomendados por el modelo de Rasch. Es decir, todas las 33 escalas creadas mantuvieron un puntaje inferior a 2 en el Eigenvalue como resultado al Análisis de componentes principales de los residuos, lo cual implica que las estas escalas, a nivel individual, no están midiendo más de un constructo a la vez (Prieto y Delgado, 2003).

En lo que respecta a las diferencias por género, solamente el dominio de Apego mostró la existencia de diferencias significativas, esto según los resultados obtenidos al contrastar las medias de ambos grupos en un análisis T de Student. De esta manera, parece que hombres y mujeres tienden a manifestar características distintas en este tipo de rasgos desadaptativos, y no es producto del azar.

No obstante, ningún otro dominio presentó diferencias significativas debido al género de los participantes. Esto puede deberse a que durante la validación de las escalas se utilizaron criterios estrictos para eliminar cualquier funcionamiento diferencial de los ítems, logrando disminuir la presencia de respuestas sesgadas a este nivel. Con base en estos hallazgos, se puede concluir que el instrumento construido funciona adecuadamente y la variable género de los participantes no presenta un posible sesgo para la medición de los distintos dominios de psicopatía.

En cuanto a las comparaciones del nivel medio de rasgos psicopáticos en distintos grupos de edad, los datos encontrados fueron de mucho valor para esta investigación, ya que en ellos se muestra una disminución significativa en la prevalencia de diferentes dominios del constructo de psicopatía, específicamente los de Conductual, Apego, Dominancia y Sí Mismo, a lo largo de los años de vivencia de las personas. Siendo los contrastes más notorios entre los grupos de 18 a 30 años y el grupo de 46 años y más. A excepción del dominio de Apego, en el cual el grupo de 31 a 45 años difirió igualmente con el grupo de 46 en adelante.

Este tipo de disminución en la manifestación de rasgos psicopáticos es algo que ya ha sido estudiado anteriormente, y que ha arrojado resultados muy similares. Tal es el caso de (Vachon et al., 2013) donde se observó una disminución en la prevalencia de rasgos psicopáticos, de 35-30%, en un grupo de edad de entre 16 y 30 años, en comparación un grupo de edad de 46 a 70 años, donde la prevalencia de estos rasgos oscilaba el 20%. De igual manera Putkonen et al., (2010), realizaron una investigación con dos grupos de edad, el primero con una media de edad de 66.8, siendo la edad mínima para este grupo 60 años, y el segundo grupo con una media de 34.8.

Estos grupos difirieron significativamente en sus puntajes de evaluación del PLC-R, siendo que el grupo de mayor edad, presento un puntaje significativamente menor en promedio (9.2), que el grupo comparativo (17.6). Interesantemente, estos grupos no difirieron en el Factor 1 del PCL-R, pero si lo hicieron significativamente en el Factor 2, donde el grupo de adultos mayores un puntaje promedio de 4.1, mientras que el grupo joven de 10.2. Esto es interesante ya que el Factor 2 evalúa rasgos y situaciones que están ligadas a los dominios en los cuales se vieron diferencias significativas en los ANOVAS.

Es evidente la relación del Factor 2 con el dominio Conductual, ya que la mayoría de sus componentes son de esta índole, pero también se ve una correspondencia con los dominios de Apego, Dominancia y de Sí Mismo, con elementos como Estilo de Vida Parasitario, en el cual escalas como Indiferente, Sin Empatía, y No Adquiere Compromisos del dominio de Apego son componentes importantes para la presencia de este tipo de rasgos. Así como también lo es lo evaluado por las escalas de Manipulador, Insincero y Charlatán en el dominio de Dominancia, y de Autojustificación y Sentido de Derecho en el dominio de Sí Mismo.

Esta información se suma a la variedad de datos que ha arrojado este instrumento, comprobando no solo su funcionalidad para detectar los rasgos deseados, sino para replicar modelos teóricos de psicopatía que ya han sido altamente estudiados. En este caso, siendo la prueba un buen parámetro para detectar diferencias entre manifestaciones psicopáticas, a lo largo del continuo de edad de las personas.

En lo que respecta a la validez de criterio, se obtuvieron resultados de mucha relevancia para el estudio. Particularmente en los análisis de correlación con el MMPI-2, es interesante ver como la escala Antisocial (Mulay et al., 2019), escala basada en el modelo alternativo de los trastornos de personalidad del DSM-5, correlacionó con todos los dominios del instrumento construido; esta escala se define por dificultades en las áreas de identidad, autodirección, empatía, intimidad, así como los rasgos patológicos de antagonismo y desinhibición, como son manipulación, insensibilidad, engaño, hostilidad, asunción de riesgos, impulsividad e irresponsabilidad (American Psychiatric Association, 2013), por lo que se convierte en una escala de gran interés para abordar en comparación con el instrumento construido. De esta forma, se aprecia cómo estas características coinciden con síntomas del CAPP ubicados a lo largo de las seis dimensiones, lo que explica las asociaciones lineales de moderadas a altas de todos los dominios con esta escala Antisocial.

Similarmente, las escalas de PAS, DIS y RC4 se correlacionan con todos los dominios a excepción de Sí Mismo. Igualmente, caracterizadas por rasgos y síntomas circunscritos a lo largo de los seis dominios del CAPP (Nichols et al., 2011); donde DIS, al ser más conductual, presenta correlaciones moderadas, pero más altas en los aspectos conductuales con los dominios Dominancia y Conductual. De acuerdo con Friedman et al. (2015), puntajes altos en la escala DIS están relacionados con una estructura de personalidad desinhibida, falta de regulación del impulso, por lo que concuerda con las correlaciones altas en cuestiones conductuales y cognitivas de nuestro instrumento.

RC4 presenta una correlación muy alta con el dominio Conductual, y moderado con Dominante, por la misma razón que DIS; sin embargo, al presentar un contenido interpersonal reflejado en malas relaciones por problemas conductuales (Friedman et al., 2015), también correlaciona moderadamente con los dominios Emocional, Apego y Cognitivo con base en escalas como falta de estabilidad emocional, Sin empatía, Indiferente e Intolerante por el síntoma de hipercrítico.

Respecto a la escala PAS, al presentar rasgos emocionales interpersonales y cognitivos (Nichols et al., 2011), correlaciona más alto con los dominios de Apego y Cognitivo, además del Conductual, pero también correlaciona moderado alto con Emocional y Dominancia. La alta correlación con el dominio de Apego puede ser explicada por el componente de Falta de remordimiento y síntomas como Sin vergüenza, Sin disculpas y Sin arrepentimiento, y su contraparte de la escala PAS. La correlación con el aspecto cognitivo puede radicar en el control de impulsos, mientras que el aspecto conductual por estos mismos comportamientos antisociales.

De acuerdo con Friedman et al. (2015), puntuaciones elevadas en la escala RC9, caracterizan individuos irritables y hostiles, además de búsqueda de sensaciones y poco control de impulsos. De esta manera, se explica la alta correlación con los dominios Dominancia y Conductual; asimismo, la correlación con el dominio Cognitivo tiene mucho que ver con los rasgos de control impulsivo. Interesantemente, esta escala se correlaciona con el dominio de Sí Mismo, lo que puede deberse a en gran medida al contenido de las escalas de Sentido de Unicidad y Sentido de Invulnerabilidad.

Otra prueba con la que se obtuvieron evidencias de validez de criterio fue el PAI; de las 15 escalas y subescalas, solamente nueve presentaron correlaciones entre moderadas y

altas. Al igual que la escala Antisocial del MMPI-2, la escala ANT del PAI mostró correlaciones con los seis dominios del instrumento construido. De acuerdo con Morey (2011), esta escala “proporciona una evaluación de características de personalidad y comportamentales relevantes a los constructos de personalidad antisocial y psicopatía” (p. 73); de manera que una moderada y alta correlación con los seis dominios del instrumento construido, brinda valiosas evidencias de correlación y validez de criterio.

Puntuaciones altas en esta escala indican irresponsabilidad, poco éxito en sus aspectos sociales y laborales, insensibilidad, relaciones frías e instrumentales, comportamiento impulsivo, temerario y hostil. Puntuaciones más bajas, pero igualmente elevadas reflejan egocentrismo, escepticismo, desinhibición y poca profundidad emocional (Morey, 2011). Con base en estas características, podemos esperar correlaciones altas a lo largo de la dimensión del constructo psicopático capturada en los seis dominios del instrumento.

Esta escala de Rasgos antisociales, a su vez, se divide en tres subescalas. ANT-A (conductas antisociales), al presentar un componente mayormente conductual, correlaciona, de manera esperable con los dominios Conductual y Dominancia. ANT-E (Egocentrismo), por su parte, tiene correlaciones moderadas con los dominios Sí Mismo, Dominancia, Conductual, y alta con Cognitivo, lo que se puede explicar por los síntomas de la escala Suspica e Intolerante. Búsqueda de sensaciones (ANT-B), al igual que escalas de desinhibición y falta de control de impulsos, correlacionó con los dominios Conductual y Cognitivo, lo que refleja estos síntomas de falta de regulación conductual y análisis de consecuencias de dichos comportamientos.

Similarmente a la escala ANT-B, la escala Desinhibición correlacionó alto con estos dominios, podemos concluir que por las mismas razones. No obstante, e interesantemente, la escala presentó una correlación moderada con el dominio Emocional; esto podría explicarse por cargas factoriales leves de ítems (del PAI) con rasgos de inestabilidad afectiva, egocentricidad, relaciones negativas y despreocupación (Ruiz et al., 2018), rasgos presentes en las escalas Falta de profundidad emocional y Falta de ansiedad del dominio Emocional.

MAN es una escala que representa síntomas de manía e hipomanía, elevación anormal del estado de ánimo, expansividad, grandiosidad, hiperactividad, irritabilidad e impaciencia (Morey, 2011). Estos síntomas parecen explicar la correlación alta que tuvo con el dominio de Sí Mismo, donde se encuentran síntomas como Sentido de ser extraordinario y Sentido de

ser merecedor en las escalas de Autoagrandecimiento, Sentido de invulnerabilidad, Sentido de derecho y Sentido de invulnerabilidad. Además del componente conductual propio de estados de manía e hipomanía, representado de la correlación con los dominios Dominante y Conductual.

MAN-A es una subescala de MAN, Nivel de actividad, que evalúa niveles de actividad ideativos (como fuga de ideas) y conductuales (como la actividad motora) (Morey, 2011). Esto se evidencia en las correlaciones con los dominios Sí Mismo, Dominancia, Conductual y Cognitivo, al presentar rasgos de sobreactivación conductual, falta de autogestión y falta de planeamiento.

La última escala del PAI que correlacionó significativamente fue AGR-F (Agresiones físicas), subescala de AGR, que correlacionó solamente con el dominio Conductual. Esta subescala, evalúa rasgos agresivos y de violencia física; más allá de los componentes conductuales que se pueden ver reflejados en las correlaciones con los dominios Conductual y Dominancia, también las correlaciones moderada y alta con los dominios Emocional y Cognitivo, respectivamente, pueden reflejarse en los síntomas de falta de empatía y falta de regulación en el accionar de conductas violentas.

En lo que respecta a las correlaciones con el UPPS-P, con la excepción del dominio de Apego, la totalidad de esta prueba muestra una correlación significativa con los cinco dominios del instrumento construido. Siendo el Conductual una muy alta correlación, seguido por el Cognitivo con una alta y luego moderadas para los demás dominios. Al igual que en la correlación con la mayoría de las escalas de los otros instrumentos, se aprecia una mayor correlación en aspectos conductuales, lo que puede deberse a un mayor contenido conductual en los ítems y escalas de estos instrumentos.

Específicamente, la escala de Urgencia positiva correlacionó alto, también, con el dominio Conductual, lo cual era algo esperado al ser está definida por una tendencia a perder el control del comportamiento ante emociones positivas intensas (Gray et al., 2019). No obstante, más allá del aspecto conductual, esta descripción también implica componentes de autocontrol y de pensamiento, al correlacionar moderadamente con los dominios de Sí Mismo y Cognitivo, se evidencia esta relación. Que puede estar fundamentada en escalas como Autoconcepto inestable, con un sentido caótico de ser, Sentido de invulnerabilidad, Sentido de ser único, con un componente de necesidad de atención, y Falta de planeamiento.

Otra correlación interesante fue la de Urgencia negativa, presentando un valor alto con el dominio Conductual y uno moderado alto con el dominio Cognitivo. Puntuaciones altas en esta escala indican una falta de control comportamental ante emociones negativas intensas y la ignorancia de las consecuencias que estas le pueden traer (Gray et. al., 2019). Por esto, y similarmente que con la escala DIS del MMPI, y Urgencia positiva, estas correlaciones tienen que ver con los aspectos conductuales y cognitivos relacionados con desinhibición del comportamiento, falta de autocontrol y planeamiento.

Con respecto a las asociaciones encontradas con el IDPP, es importante comprender por qué la escala de psicoticismo fue la única en correlacionar altamente con todos los dominios del instrumento. Orozco y Quesada (2018), explican que el psicoticismo está compuesto por cuatro factores: 1) Creencias y experiencias inusuales, 2) Excentricidad, 3) Desregulación cognitiva y perceptual, y 4) Susplicacia y Paranoia.

En el caso de las Creencias y Experiencias Inusuales, Krueger y Markon (2014) explican cómo esta puede estar ligada a la creencia de poseer habilidades inusuales y excepcionales, rasgo que comparte relación con las escalas de Sentido de Unicidad, que está formado por los síntomas de Sentido de ser Extraordinario, Excepcional y Especial, y con la escala de Sentido de Invulnerabilidad, que está formada por los síntomas de Sentido de Ser Invencible, Indestructible, Imbatible, con lo cual, era esperable una alta correlación con el Dominio de Sí Mismo, a la cual pertenecen las escalas antes mencionadas.

Adicionalmente, en este mismo apartado de Creencias y Experiencias Inusuales, puede estar relacionado altamente con creencias religiosas y experiencias espirituales (Wolfradt et al., 1999), con lo cual, se vuelve evidente la existencia de una correlación con el dominio Cognitivo, mismo que en la Escala de Intolerante, incluye el síntoma de Fanático, el cual puede manifestarse con un alto componente religioso.

Por otro lado, el factor de la Excentricidad está ligado a exhibir comportamientos, habla peculiares, inusuales y extravagantes (Krueger y Markon, 2014), estos rasgos están altamente ligados a dominio de Sí Mismo, ya que este dentro de sus escalas contiene la de Egocéntrico, mismo que evalúa el síntoma de Presuntuoso, que tiene una tendencia a mostrar comportamientos extravagantes. De igual forma, existe una correlación con el dominio de Dominancia, mismo que tiene la escala de Charlatán, esta escala evalúa el síntoma de Pretensioso, que a su vez está ligado con la exhibición de comportamientos antes expuesta.

Adicionalmente, Esbec y Echeberúa (2011), mencionan que personas con alta Excentricidad tienden a presentar una capacidad reducida para las relaciones interpersonales, es por ello, que también era esperable su correlación con el dominio de Apego, especialmente, con la escala de Desapego, dado a su síntoma de Reservado.

A lo que respecta el factor de Desregulación Cognitiva y Perceptual, Krueger y Markon (2014) indican que este puede estar ligado a aspectos de despersonalización, desrealización, y experiencias disociativas. Con lo cual esperábamos una correlación importante con el dominio Emocional, mismo que posee la escala de Falta de Estabilidad Emocional. Finalmente, con el factor de Suspiciousidad o Paranoia, era predecible una correlación con el dominio Cognitivo, el cual contiene toda una escala de Suspiciousidad para evaluar dicho constructo, por lo que no es de extrañar la alta correlación obtenida.

Este mismo escenario lo podemos ver reflejado con la alta correlación existente entre el factor de Desapego, y el dominio de Apego del instrumento creado, esta relación era esperable dado que el dominio de Apego contiene una escala para evaluar el Desapego, con lo cual esta alta correlación es un excelente indicador de la eficacia de la escala creada.

Otro de los aspectos importantes a destacar son las altas correlaciones que obtuvo el factor de Antagonismo del IDPP, especialmente con el dominio de Dominancia. Esto debido a que Orozco y Quesada (2018) destacan en el antagonismo la existencia de rasgos como la Manipulación y el Engaño. Dentro del instrumento creado, existen una escala para evaluar la Manipulación, y otra para evaluar el Engaño, ambas pertenecientes a la Dominancia, con lo cual dicha correlación es de mucho valor para los hallazgos de este estudio.

En cuanto al factor de Desinhibición, era de mucho interés para esta investigación ver su correlación con el dominio Conductual, esto debido a que Orozco y Quesada (2018), incluyen dentro de dicho factor, rasgos como lo son la Irresponsabilidad e Impulsividad, dichos rasgos son evaluados en este instrumento bajo las escalas de Poco Confiable, la cual incluye irresponsabilidad, y en la escala de Imprudente, la cual incluye precipitado y que viene a evaluar lo mismo que la impulsividad. La correlación entre dicho factor y el dominio Conductual, no solo fue muy alta como se esperaba, si no que fue la más alta entre todas las correlaciones entre los dos instrumentos, llegando a una puntuación de $r=.763$.

Por otro lado, en cuanto a las correlaciones vistas en el NEO PI-R, inicialmente se destaca la del Neuroticismo. Los puntajes elevados de neuroticismo en el NEO PI-R, suelen ser un indicador de que la persona evaluada es propensa a experimentar angustia, ideas poco realistas, impulsos excesivos y respuestas de afrontamiento desadaptativas (Costa y McCrae, 2008). Es por ello que, al realizar la correlación de Pearson, esperábamos una correlación significativa con los dominios Conductuales y Cognitivos, expectativa que fue cumplida.

En el caso de dominio Conductual, la correlación alta puede ser explicada debido a la presencia de escalas de Agresividad, Imprudencia e Intranquilidad, factores que juegan un papel importante en respuestas de afrontamiento desadaptativas. Mientras que, en el dominio Cognitivo, el cual presentó una correlación moderada, podemos encontrar una relación con el sentido de angustia dado que este dominio presenta la escala de Suspica, misma que evalúa síntomas como la hipervigilancia, cautela y desconfianza. A su vez, este mismo dominio podría presentar relación con la posible existencia de ideas poco realistas del neuroticismo, debido a que la escala de Intolerante busca detectar síntomas de hipercrítico, pensamiento estrecho y fanatismo.

Adicionalmente, uno de los puntos más cruciales por el cual se tomó la decisión de correlacionar con el NEO PI-R, fue la posibilidad de poner a prueba la validez convergente de manera directa. Especialmente con los factores de Responsabilidad y Amabilidad. En el caso de la Responsabilidad, encontramos una correlación negativa significativa con todos los dominios del instrumento, de estas correlaciones, sólo la que obtuvo con el dominio de Sí Mismo fue baja, las restantes fueron como mínimo moderadas. Sin embargo, la que destaca y resulta de más importancia es la correlación con el dominio Conductual. Esta correlación fue alta, y tiene especial valor ya que este dominio tiene una escala de Poco Confiable, que tal y como se expresó anteriormente, busca detectar rasgos de irresponsabilidad, con lo cual era esperable e importante que se diera esta correlación negativa, ya que es un indicativo claro de la funcionalidad de la escala para detectar estos rasgos.

Por último, con el factor de la Amabilidad, Costa y McCrae (2008) explican que este define interacciones por las cuales una persona tiene preferencias, yendo desde la compasión hasta el antagonismo en un continuo. Esto es crucial ya que pone a este factor en un encuadre dimensional, mismo que nosotros buscamos obtener con esta investigación. Adicionalmente, estos autores explican que un puntaje alto implicaría rasgos como bondad, confianza, y una

anuencia a ser serviciales. Mientras que un puntaje bajo se manifestaría en personas cínicas, groseras, desconfiadas, irritables y hasta manipuladoras.

El factor de Amabilidad obtuvo correlaciones moderadas con todos los dominios, siendo los de menor puntaje el dominio Conductual y el Cognitivo, contrario a otras correlaciones que ya hemos explicado. Este hallazgo es positivo para la investigación, ya que estos dominios son los que menos rasgos comparten con lo descrito por Costa y McCrae (2008) para el factor de amabilidad, lo cual refuerza el hecho de que las correlaciones obtenidas por esta prueba no son aleatorias.

En cuanto al dominio de Sí Mismo y el dominio de Apego, la correlación negativa se entiende debido a su escala de Sentido de Derecho y Sentido de Desapegado respectivamente. Las cuales contemplan rasgos como la exigencia y frialdad, mismos que pueden ser encontrados en una persona cínica. Por su parte, el dominio Emocional se esperaba una correlación negativa ya que, en su escala de Falta de Estabilidad Emocional, uno de los síntomas evaluados es la irritabilidad, la cual como ya se mencionó anteriormente, es uno de los rasgos vistos en personas con puntajes bajos en la escala de Amabilidad del NEO PI-R (Costa y McCrae, 2008). Finalmente, del dominio de Dominancia, se podía esperar una correlación negativa alta, la cual estuvo muy cerca de suceder (el puntaje fue de $r=-.598$), esto se explica debido a que en ella se encuentra una escala que mide el rasgo de manipulación, además de otras como Insincero y Engañoso, que están relacionadas con este constructo. De esta manera, podemos tomar esta correlación negativa como una garantía del funcionamiento del dominio creado.

Como se ha podido apreciar, el dominio Conductual presenta correlaciones con varias de las escalas con las que se comparó. Esto puede deberse a que las conductas son una forma más evidente de evaluar síntomas, al presentar los rasgos de una manera más tangible y observable. Además, de ser la forma más sencilla para las personas de identificar y reportar los temas abordados en el contenido de los ítems.

Otro punto interesante de mencionar es que las escalas que evalúan el constructo psicopático a un nivel más general, como Antisocial del MMPI-2, ANT del PAI y los dominios de Antagonismo y Desinhibición del IDPP, correlacionan altamente con los seis dominios planteados por el CAPP y por consiguiente de nuestro instrumento. Asimismo, las correlaciones presentadas por el NEO PI-R, especialmente los factores de Amabilidad y

Responsabilidad muestran patrones de asociación significativos y altos con los dominios correspondientes de instrumento de psicopatía construido en esta investigación.

Sumado a esto, está el hecho que el dominio Conductual no correlacionó con cualquier escala, lo cual pudo haber sido el caso, ya que como se expuso anteriormente, las conductas tienden a correlacionar con mayor facilidad por su naturaleza observable; correlacionó significativamente, pero muy bajo, con la escala de Desapego del IDPP ($r=.224$), por ejemplo, lo que se repitió en muchos otros de los indicadores que no han sido incluidos en los resultados de este estudio, por una cuestión de espacio. Si bien es cierto reportamos evidencias de validez convergente, también se obtuvieron resultados satisfactorios en relación con la validez discriminante.

Por tanto, concluimos que lo anterior aporta evidencias positivas para nuestro modelo, ya que las escalas que evalúan el constructo a un nivel más amplio correlacionan con los diferentes dominios planteados a lo largo de la dimensión, lo que demuestra la capacidad de nuestro instrumento de captar diferentes síntomas a lo largo del continuo del constructo. Asimismo, podemos afirmar que los dominios del instrumento creado muestran una capacidad de ajuste destacable con respecto a las escalas de los otros instrumentos ampliamente validados e investigados, de los cuales se esperaba una correlación significativa, siendo que en todos los casos se encontraron correlaciones, no sólo estadísticamente significativas, sino que con valores entre moderados a altos. Esto mismo se ve respaldado por los datos obtenidos de análisis factorial exploratorio.

Inicialmente se hipotetizó que el modelo de psicopatía presentaría 6 factores, uno por cada dominio; sin embargo, el análisis factorial brindó evidencias de un modelo de psicopatía basado en tres factores. Si bien, algunas escalas cargan en factores distintos a los que se esperaba que lo hicieran, y algunas, presentan solapamiento con otros factores, el modelo resultante es consistente con las investigaciones más modernas del modelo del CAPP (Sellbom et al., 2021).

Para el instrumento construido, el factor 1 se compone principalmente de escalas conductuales y de falta de habilidades, propias de los seis dominios, pero principalmente el Conductual, Emocional y Cognitivo; esto refleja un contenido de desregulación conductual. A su vez, el factor 2 está compuesto por síntomas que tienen que ver con la forma de ser, escalas principalmente del dominio de Sí Mismo, pero también escalas del dominio

Dominante y Cognitivo que refieren a rasgos y síntomas propios de la forma de ser de una persona, con tendencia a la dominancia, grandiosidad y narcisismo. Finalmente, el factor 3 parece estar relacionado con un factor afectivo y emocional, compuesto principalmente por escalas del dominio de Apego, Emocional, y Cognitivo, reflejando deficiencias en el procesamiento emocional relacionado al funcionamiento intra e interpersonal.

De esta manera, las escalas cargaron sus pesos factoriales de manera muy similar a las agrupaciones factoriales reportadas por Sellbom et al. (2021). Ellos determinaron un modelo de tres factores para el instrumento de autorreporte CAPP-SR, estos factores fueron Antagonismo/Mezquindad (Factor 1), Desinhibición (Factor 2) y Grandiosidad Sin Miedo (Factor 3).

En comparativa, el Factor 1 del instrumento construido es muy similar al Factor 2 del CAPP, Desinhibición, con excepción de las escalas de Desapegado, Engañoso, Manipulador, Centrado en Sí Mismo y Sin Capacidad para Experimentar Placer, que cargaron en otros factores. Por su parte, el Factor 2 es similar al factor Grandiosidad Sin Miedo donde, del instrumento construido, faltó solamente la escala Inflexible, y cargaron adicionalmente las escalas Engañoso, Manipulador, Indiferente y Sin Capacidad para Experimentar Placer. Por último, el Factor 3 se asemeja al factor Antagonismo/Mezquindad, donde la mayoría de las escalas coinciden, y solo destaca la ausencia de Falta de Ansiedad, y se adicionaron las escalas de Inflexible y Centrado en Sí Mismo, por parte de nuestro instrumento. Debido a estas similitudes, y a pesar de las diferencias, la estructura general de las cargas factoriales, fue lo suficientemente similar a las estructuras propuestas por investigación previa bajo este modelo del constructo psicopático.

Al igual que la similitud con los factores presentados por el modelo CAPP (Sellbom et al., 2021), los síntomas o rasgos de las escalas construidas en el presente trabajo, según su carga factorial, tienen similitud con factores planteados por otros modelos explicativos del constructo psicopático. Por ejemplo, el factor 1, se compone de escalas de síntomas y rasgos más conductuales, impulsivos y con tendencias a quebrantar reglas, síntomas que se reflejan en el clásico Factor 2 del PCL-R, Estilo de vida antisocial (Harpur et al., 1988, como se citó en Sellbom et al., 2021). Asimismo, este factor refleja el factor de Impulsividad y comportamiento irresponsable del modelo del PCL planteado por Cooke y Michie (2001, como se citó en Sellbom et al., 2021), y del constructo de desinhibición concebido por el

modelo Triárquico y la medida del EPA (Few et al., 2013; Lynam et al., 2013, como se citaron en Sellbom et al., 2021).

El factor 2, por su parte, al representar síntomas relacionados a una forma de ser dominante, grandiosa, narcisista y de falta de miedo, se asemeja con elementos de dominancia y falta de miedo presentes en el PPI y de audacia del modelo Triárquico (Lilienfeld y Andrews, 1996; Benning et al., 2003; Patrick et al., 2009, como se citó en Sellbom et al., 2021). Similarmente, este factor se alinea con el factor de Estilo interpersonal engañoso del modelo de 3 factores del PCL (Cooke y Michie, 2001, como se citó en Sellbom et al., 2021), al presentar rasgos de charlatanería, manipulación y grandiosidad.

Muchos de los síntomas identificados al factor 3, afectivos y de funcionamiento intra e interpersonal, tienen similitud con el Factor 1 (Afectivo-Interpersonal) del PCL-R (Harpur et al., 1988, como se citó en Sellbom et al., 2021), así como con el factor de Experiencia Afectiva Deficiente del modelo de Cooke y Michie (2001, como se citó en Sellbom et al., 2021). Además, este primer factor, por su contenido, es consistente con el dominio de mezquindad planteado por el modelo triárquico de psicopatía (Patrick et al., 2009, como se citó en Sellbom et al., 2021) y el domino antagonista planteado por el modelo de medición del EPA (Few et al., 2013; Lynam et al., 2013, como se citaron en Sellbom et al., 2021).

Estos hallazgos expuestos a lo largo de la discusión son de enorme valor para este estudio, ya que con ellos se puede concluir que el instrumento creado demuestra una alta funcionalidad a lo largo de sus dominios y gran mayoría de sus escalas, y además ha logrado ajustarse en sus tres factores, no solo a las investigaciones más modernas del modelo teórico bajo el cual se abordó, el CAPP, sino que también logró establecer sus factores con un alto nivel de consistencia con los modelos de psicopatía más respetados y estudiados a nivel mundial.

Esto a su vez, implica un avance enorme en la investigación de la psicopatía, no solo en Costa Rica, sino a nivel latinoamericano, donde los estudios de este tema son reducidos y no han existido, hasta el momento, intenciones de crear pruebas psicométricas para su detección, más allá de adaptar pruebas ya existentes en otros países y continentes.

Los resultados de este instrumento podrán ser utilizados en futuras investigaciones del tema como una base sólida de la cuál partir, ya que los datos obtenidos muestran un alto

potencial por su eficiencia, confiabilidad, y ajuste con modelos teóricos e instrumentos existentes, lo cual toma especial valor si se toma en consideración el hecho de que este es apenas el primer corte de resultados de esta prueba psicométrica.

Conclusiones

Al final de esta extensa investigación, podemos confirmar que la mayoría de los objetivos han sido cumplidos a cabalidad, mientras que otros fueron afectados por las limitaciones que se expondrán en el apartado siguiente. Sin embargo, en el caso de los objetivos no culminados, se hicieron avances importantes para su eventual finalización.

Con respecto al primero de los objetivos generales, se logró construir una prueba psicométrica auto aplicada que logra capturar rasgos psicopáticos en la población costarricense, y en la cual, las escalas logran diferenciar distintos niveles de severidad de estos rasgos en la persona, con lo cual se consigue también que esta prueba psicométrica cumpla con la perspectiva dimensional propuesta.

Esta prueba psicométrica para evaluar rasgos psicopáticos de forma autoaplicada, es la primera que se ha creado con esta finalidad en el país, donde no hubo restricciones con base a la provincia de residencia durante su creación, y se obtuvieron datos de personas de todos los niveles de escolaridad, lo cual es fundamental para entender su eficacia en futuras personas que puedan ser evaluadas con este instrumento.

Por otra parte, se logró que los enunciados transmitieran de forma clara y concisa las ideas principales que se querían abordar, sin inducir a una opción de respuesta a la persona participante. De igual manera, se logró evitar que los ítems se cargaran de información irrelevante que pudieran sesgar las respuestas de las personas, que no existieran oraciones complejas o de difícil comprensión, y se redujo considerablemente el tiempo de aplicación del instrumento eliminando los ítems de ajuste inadecuado.

Adicionalmente, con el segundo objetivo general, podemos afirmar que no solo se construyó este instrumento de evaluación psicométrico, si no que se sometió a diversos análisis estadísticos con la finalidad de garantizar su eficacia y ajuste apropiado. Siguiendo el modelo de Rasch, se hicieron análisis de Infit y Outfit para comprender el ajuste tanto de los ítems como de las personas hacia el instrumento. Se sometió a revisión el nivel de fiabilidad

y separación tanto de ítems como de la muestra, y se descartaron ítems que generaban sesgo en los resultados a través del análisis DIF, además se hizo un Análisis de Componentes Principales de los residuos para determinar la unidimensionalidad de las escalas. Todo lo anterior sumado a análisis más convencionales y altamente utilizados como lo son el Alfa de Cronbach, prueba T Student, Análisis Factorial Exploratorio, y ANOVA.

Gracias a esta variedad de análisis a la que se sometió la prueba, se logró aportar una robusta cantidad de datos que refleja y demuestra que hasta en las escalas que presentaron resultados moderados, se puede afirmar que existe suficiente información para considerarlas como una base fuerte y avanzada para terminar de confirmar la suficiencia de estas. Dicho esto, y de la mano con lo propuesto en el objetivo específico tres y con los resultados de esta investigación, queda comprobado, que una importante mayoría de las escalas creadas mostró sensibilidad para detectar los diferentes niveles de respuesta de las personas, un ajuste apropiado a la muestra, datos de confiabilidad y consistencia interna fuertes, y en todos los casos, se logró establecer la unidimensionalidad del constructo en cuestión.

Estos logros se pueden adjudicar a la cuidadosa planificación que se tomó desde el inicio para la creación de instrumento, donde se hizo un arduo esfuerzo porque cada una de las escalas tuvieran ítems que evaluaran conductas, pensamientos, emociones y actitudes. De esta manera se evitó que dominios con componentes normalmente conductuales, se vieran fortalecidos por sobre otras escalas, y a su vez, fue un insumo a largo plazo, para detectar los niveles de severidad a lo largo de la dimensionalidad de los constructos. Esto a su vez cumplió una doble función, ya que también logró abordar de forma efectiva el primer objetivo específico. Logrando que el constructo de psicopatía planteado detrás del instrumento sea de múltiples factores, tres específicamente, esto a su vez, se alinea con lo visto en los estudios más recientes del modelo del CAPP (Sellbom, et al., 2021).

Además, en relación con el cuarto objetivo específico, y el objetivo externo, podemos confirmar que se dieron pasos en función de desarrollar normas de interpretación del instrumento, de hecho, se hizo una guía de interpretación rápida (ver anexo 6), con la cual se permite hacer conclusiones iniciales de las respuestas obtenidas. La razón por la cual no se ahondó en estas normas o en un manual de interpretación robusto, es debido a que este instrumento requiere estudios posteriores y algunas modificaciones puntuales, pero importantes que detallaremos en las recomendaciones. Con lo cual, es pertinente hacer el

desarrollo de dichas normas y manual, una vez que se esté en dicha etapa de la investigación de este instrumento, misma que continuará posterior a la finalización de este trabajo final de graduación.

Adicionalmente, es importante mencionar que los datos obtenidos y expuestos son con la población normativa costarricense. Obtener información estadística de la efectividad de esta prueba en población forense era parte de algunos de los objetivos planteados en esta investigación; no obstante, más allá de que se hicieron esfuerzos por obtener acceso a la población forense del país, debido a circunstancias ajenas, relacionadas a la pandemia, esto no fue posible, y se explicará con mayor detalle en el apartado de limitaciones. Dicho esto, después de muchos esfuerzos intentando acceder a esta población, se logró recopilar los resultados de 18 personas privadas de libertad. Pero desafortunadamente, estas no fueron tomadas en cuenta para hacer análisis, ya que la muestra fue considerada muy pequeña para arrojar resultados útiles.

Por concluir y alineado a nuestros objetivos, el presente trabajo presenta la construcción de una prueba de autorreporte para evaluar el constructo psicopático, lo cual representa ciertos obstáculos, tal y como lo plantean Sellbom et al. (2018). A pesar de estos cuestionamientos, la construcción de esta prueba se realizó tomándolos en cuenta, con el fin de evitar sesgos y deficiencias que arriesgaran su confiabilidad y validez, y así garantizar una medición efectiva del constructo de psicopatía de forma dimensional, cumpliendo lo propuesto por el modelo CAPP, una medida de rasgos dimensionales que existen en cualquiera y mayormente a un nivel no patológico, por lo cual es apropiado para medir la psicopatía a través de diferentes contextos (Sellbom, Cooke y Shou, 2019).

Limitaciones

La principal limitación del presente trabajo fue el surgimiento del virus SARS-CoV-2 y la pandemia que provocó alrededor de la enfermedad por coronavirus (Covid-19). Las implicaciones de la pandemia, la cuarentena, cierre de instituciones, imposibilidad de acceso a las personas frente a frente, se convirtieron en grandes obstáculos para las diferentes etapas y tareas a realizar en la elaboración de este trabajo.

Uno de los obstáculos más retadores y limitantes por la pandemia fue el cierre de acceso a los centros penitenciarios, y, por ende, el acceso a la población forense. Debido a

esto se vio imposibilitada la recolección de esta muestra. Esto tuvo un impacto muy significativo para el trabajo, ya que el no contar con una muestra forense, afecta directamente los objetivos.

Pruebas como el PCL-R, que requieren acceso a información corroborativa, implican obstáculos y mayor dificultad para investigación o abordaje en poblaciones no institucionalizadas donde no existe un expediente en el que se pueda consultar dicha información (Sellbom, et al., 2018). De esta manera, el imposibilitado acceso a la población forense tuvo como consecuencia adicional el no poder comparar las mediciones del instrumento construido con “la regla de oro” de la medición del trastorno de personalidad psicopático, el PCL-R.

No obstante, con la colaboración de psicólogos forenses del Centro de Atención Integral Jorge Arturo Montero Castro, se logró la aplicación de la prueba construida a algunos participantes de esta población forense. A pesar de estos esfuerzos, como se explicó en el apartado anterior, la cantidad de personas evaluadas de muestra forense no fue suficiente para un análisis psicométrico confiable y robusto; razón por la cual se toma la decisión de descartar esos resultados.

Adicionalmente, a pesar de la anuencia y colaboración del equipo de psicólogos forenses, por razones logísticas y las limitaciones dadas por la pandemia, no se contó con un acercamiento a las personas que administrarían la prueba, para establecer una relación más amplia que consideramos necesaria para el abordaje y estudio de la población forense de una manera óptima. Particularmente considerando las limitaciones planteadas por Sellbom et al., (2018), donde la deseabilidad social, rasgos, deshonestos y posible mentira instrumental o patológica, son variables que pueden introducir un posible sesgo en la evaluación; teniendo en cuenta que, para estas personas evaluadas, la aplicación de la prueba vino de mano del propio personal de la institución, la posibilidad de sesgo por querer dar una impresión positiva es alta.

Similarmente, el acceso a la población general se vio restringido por la cuarentena. La posibilidad de entrevistar y evaluar a las personas en un contexto presencial implicaba un riesgo significativo, por lo que se tuvo que recurrir a la virtualidad vía medios digitales como lo son cuestionarios en línea, reuniones virtuales, etc.

Se pretendía realizar un muestreo estratificado mediante el cual se dividiría la población en subgrupos por sexo, edad, y tipo de delito (violentos y no violentos) en el caso de la población forense; y un muestreo por cuotas, que busca reproducir de la forma más adecuada las características de la población en la muestra (Romero y Bologna, 2011), de estos subgrupos con el que se pretendía seleccionar personas de cada subgrupo dependiendo del porcentaje que este represente de la población total. No obstante, debido a las limitaciones mencionadas anteriormente, se procedió a realizar un muestreo no probabilístico y más de conveniencia, donde se aceptó la participación de toda persona dispuesta a completar el instrumento, siempre y cuando cumpliera los criterios de inclusión.

Otra limitante significativa tiene que ver con la longitud y extensión del instrumento. La aplicación promedio tarda entre una hora y una hora y media. Debido a esto, y sumado a la aplicación virtual, el acceso y colaboración de las personas fue complicado. En cuanto a la prueba construida, en muchos casos, las personas comenzaban a completarla, pero eventualmente optaban por no concluir el instrumento. Sin embargo, para no comprometer la validez y consistencia de la prueba, se decidió no recortarla antes de obtener datos científicos que nos permitieran entender cuáles ítems debían ser eliminados.

Similarmente sucedió con las entrevistas para la aplicación del PCL: SV, y ante la negativa y poca participación de las personas a colaborar con la entrevista, se decidió descartar la aplicación de esta. Esto significó un cambio significativo, al no contar con una prueba estandarte para la medición de la psicopatía. Esta negativa también se evidenció ante la solicitud a los participantes de administrar la prueba construida una segunda vez, luego de tres meses de la primera aplicación. Debido a esto, no se obtuvieron evidencias de la estabilidad de las medidas a través del tiempo mediante la técnica de test-retest (Prieto y Delgado, 2010).

A raíz de esta limitación, es interesante abordar el sentido de obligación de las personas en la virtualidad. Se puede hipotetizar que, en el contexto virtual, las personas no le atribuyen tanta importancia a tareas o solicitudes de esta índole. Adicionalmente, se puede considerar la afectación a la salud mental de la población en el contexto de la pandemia, misma que pudo tener implicaciones con la anuencia a participar, o finalizar una vez iniciado, debido al contenido de la prueba.

Otra limitación significativa se circunscribe en la imposibilidad de acceder a una muestra no sólo más numerosa, sino mucho más heterogénea. El no poder realizar un muestreo estratificado puede generar una tendencia homogénea en la muestra, por lo que los datos pueden tener ciertas tendencias. Asimismo, el no poder incluir una muestra forense, imposibilitó el recabar datos sobre una mayor severidad en los ítems de mayor dificultad.

Finalmente, y como última limitación se puede mencionar el no poder cumplir con el objetivo externo de construir un manual de aplicación para el instrumento creado. Debido a la imperante necesidad de validar el instrumento con una muestra más heterogénea, como se mencionó en el párrafo anterior, tanto a nivel de población general como de población forense, se determinó más provechoso realizar una futura validación de la prueba y posteriormente confeccionar el manual de aplicación e interpretación con base en estos datos mayormente robustos.

Recomendaciones

Primeramente, para futuros estudios sobre el constructo de psicopatía, debido al debate conceptual existente a su alrededor, y a la investigación realizada para el presente trabajo, se recomienda el abordaje desde el marco conceptual planteado por el modelo del CAPP. Esto, con base en su caracterización dimensional de la psicopatía, que permite una evaluación de rasgos y síntomas a lo largo de todo el espectro, no solamente conductas antisociales o criminales.

Particularmente para futuras investigaciones con el instrumento construido en este trabajo final de graduación, es recomendable recabar una muestra más heterogénea, incluyendo personas de diferentes contextos, tanto en población general como forense. Utilizar un método de muestreo estratificado o probabilístico, tal y como se tenía planeado inicialmente en este estudio, podría ser de utilidad para lograr este cometido.

En esta misma línea, se recomienda la validación del instrumento con una muestra forense adecuada en cantidad de personas, para poder tener criterios confiables y realizar análisis estadísticos robustos con el fin de obtener resultados fiables sobre la eficacia del instrumento creado, en esta población.

Asimismo, a modo de recomendación, se plantea la necesidad de construir una escala de validez para el instrumento; esto con el fin de detectar y minimizar patrones de respuesta en lo contestado por las personas evaluadas y proveer así datos más confiables, eliminando sesgos como el de deseabilidad social o brindar una buena imagen por miedo a posibles implicaciones institucionales.

Además de esta escala de validez, se recomienda, para minimización de este tipo de sesgos y una vez validado el instrumento en población forense, la posibilidad de una aplicación multimétodo donde el autorreporte se complementa con otras formas de evaluación como la entrevista conductual, examen mental, observación de conducta y entrevista a personas cercanas. De esta forma se pueden triangular los datos con otras medidas y poder realizar una evaluación más completa y confiable de rasgos psicopáticos.

Referencias bibliográficas

- Abalos, C., Esquivel, N., y Gallardo, C. (2004). *Adaptación de la Escala de Calificación de la Psicopatía Revisada (PCL-R) de Robert Hare en población reclusa del Centro de Detención Preventiva de San Miguel* [Tesis de grado]. Universidad de Chile, Santiago, Chile.
- American Educational Research Association & American Psychological Association. (2014). *Standards for Educational and Psychological Testing*. Amer Educational Research Assn.
- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: APA.
- Arce, B. y Gutiérrez, C. (2003). *Psicopatía en violadores y su forma de operar en el delito: Descripción acerca de los niveles de psicopatía en los distintos tipos de violadores de mujeres y su relación en la forma de cómo cometen el delito*. [Tesis de Licenciatura]. Universidad de Costa Rica, Costa Rica.
- Bartolomé, A., & Prieto, G. (2017). Análisis del Test de Detección del Cuidador Quemado con el Modelo de Rasch. *Escritos de Psicología*, 10(2), 116–125. <https://doi.org/10.24310/espsiescpsi.v10i2.13182>
- Bond, G. T., & Fox, M. C. (2007). *Applying the Rasch model*. Lawrence Erlbaum Associates, Publishers, INC.
- Boone, W. J., & Staver, J. R. (2021). *Advances in Rasch analyses in the human sciences* (1st ed.). Springer Nature.
- Caponi, S. (2018). La psiquiatrización de la vida cotidiana: el DSM y sus dificultades. *Metatheoria*, 8(2), 97-103. doi: [20.500.11807/2514](https://doi.org/10.500.11807/2514)
- Carragher, N., Krueger, R. F., Eaton, N. R., y Slade, T. (2015). Disorders without borders: current and future directions in the meta-structure of mental disorders. *Social Psychiatry and Psychiatric Epidemiology*, 50(3), 339–350. doi: 10.1007/s00127-014-1004-z

- Cleckley, H. (1988). *The mask of sanity*. St. Louis, MO: Mosby. (Original work published 1941)
- Cooke, D. J., Hart, S. D., Logan, C. y Michie, C. (2012). Explicating the Construct of Psychopathy: Development and Validation of a Conceptual Model, the Comprehensive Assessment of Psychopathic Personality (CAPP). *International Journal of Mental Health, 11*, 242–252. doi: 10.1080/14999013.2012.746759
- Cooke, D. J. y Logan, C. (2018). Capturing Psychopathic Personality. En C. Patrick (Ed.). *Handbook of Psychopathy* (pp. 189-210). New York: The Guilford Press.
- Costa, P., y McCrae, R. (2008). *NEO PI-R. Inventario de Personalidad NEO Revisado* (3 ed.). Madrid: TEA Ediciones.
- Crego, C., y Widiger, T. A. (2014). Psychopathy and the DSM. *Journal of Personality, 83*(6), 665–677. doi: [10.1111/jopy.12115](https://doi.org/10.1111/jopy.12115)
- DeLisi, M. (2016). *Psychopathy as unified theory of crime*. New York, USA: Palgrave Macmillan.
- Echeburúa, E., Salaberría, K., y Cruz-Sáez, M. (2014). Aportaciones y Limitaciones del DSM-5 desde la Psicología Clínica. *Terapia Psicológica, 32*(1), 65–74. doi:10.4067/s0718-48082014000100007
- Esbec, E., y Echeburúa, E. (2011). La reformulación de los trastornos de la personalidad en el DSM-V. *Actas Esp Psiquiatr, 39*(1):1-11
- Esbec, E., y Echeburúa, E. (2015). El modelo híbrido de clasificación de los trastornos de la personalidad en el DSM-5: un análisis crítico. *Actas Esp Psiquiatr, 43*(5), 177-86.
- Fisher, R., & Prieto, G. (2016). Análisis del Test de Actitudes Alimentarias (EAT-26) con un modelo tipo Rasch en una muestra de adolescentes paraguayos. *Cuadernos Hispanoamericanos De, Psicología, 1*, 5–16.
- Friedman, A.F., Bolinsky, P.K., Levak, R.W., y Nichols, D.S. (2015). *Psychological Assessment with the MMPI-2/MMPI-2-RF*. Routledge.
- Gamboa, J. y Pérez, I. (2004). *Trastornos y desajustes psicológicos asociados a la violencia delictiva: Una exploración de la violencia utilizada en la comisión de los delitos por*

- parte de mujeres costarricenses privadas de libertad*. [Tesis de Licenciatura]. Universidad de Costa Rica, Costa Rica.
- Gardner, B. O., Boccaccini, M. T., y Edens, J. F. (2015). Personality Assessment Inventory Scores as Predictors of Misconduct, Recidivism, and Violence: A Meta-Analytic Review. *Psychological Assessment*, 27(2), 534-544.
- Gatner, D. T., Blanchard, A., Douglas, K. S., Lilienfeld, S. O. (2018). Psychopathy in a Multiethnic World: Investigating Multiple Measures of Psychopathy in Hispanic, African American, and Caucasian Offenders. *Assessment*, 25(2), 206-221.
- González, J., y Pazmiño, M. (2015). Cálculo e interpretación del Alfa de Cronbach para el caso de validación de la consistencia interna de un cuestionario, con dos posibles escalas tipo Likert *Revista Publicando*, 2(1), 62-77.
- Gough, H. G. (1960). Theory and method of socialization. *Journal of Consulting and Clinical Psychology*, 24, 23–30.
- Gray, N.S., Weidacker, K., & Snowden, R.J. (2019). Psychopathy and impulsivity: The relationship of psychopathy to different aspects of UPPS-P impulsivity. *Psychiatry Research*, 272, 474-482. doi: 10.1016/j.psychres.2018.12.155
- Guay, J. P., Ruscio, J., Knight, R. A., y Hare, R. D. (2007). A taxometric analysis of the latent structure of psychopathy: Evidence for dimensionality. *Journal of Abnormal Psychology*, 116, 701–716. doi: [10.1037/0021-843X.116.4.701](https://doi.org/10.1037/0021-843X.116.4.701)
- Guerra, T. (2017). Análisis factorial confirmatorio de la escala Appraisal of Self Care Agency Scale – Revise. *Rev. Latino-Am. Enfermagem*, 25, 1-9. doi: 10.1590/1518-8345.1378.2856
- Gutiérrez, F., Aluja, A., Peri, J. M., Calvo, N., Ferrer, M., Baillés, E., ... Krueger, R. F. (2015). Psychometric Properties of the Spanish PID-5 in a Clinical and a Community Sample. *Assessment*, 1-11. doi: 10.1177/1073191115606518
- Hare, R. D. (2003). *The Hare Psychopathy Checklist—Revised* (2nd ed.). Toronto, ON, Canada: Multi-Health Systems.

- Hare, R. D. (2010). PCL-R. Escala de Evaluación de Psicopatía de Hare-Revisada, Manual Técnico. Madrid, España: TEA EDICIONES.
- Hare, R. D., Neumann, C. S., y Mokros, A. (2018). The PCL-R Assessment of Psychopathy Development, Properties, Debates, and New Directions. En C. Patrick (Ed.). *Handbook of Psychopathy* (pp. 39-79). New York: The Guilford Press.
- Hart, S. D, Cox, D. N., y Hare, R. D. (1995). Escala de la Evaluación de la Psicopatía: Versión de Cribado (Adaptación al Castellano, T. Silva, E. López, Garrido, V.). Toronto, ON, Canada: Multi-Health Systems.
- Hart, S. D., Watt, K. A., & Vincent, G. M. (2002). Commentary on Seagrave and Grisso: Impressions of the state of the art. *Law and Human Behavior*, *26*(2), 241–245. doi:10.1023/a:1014648227688
- Hemphill, J. F., Templeman, R., Wong, S., & Hare, R. D. (1998). Psychopathy and crime: Recidivism and criminal careers. In *Psychopathy: Theory, research and implications for society* (pp. 375-399). Springer, Dordrecht.
- 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. A., 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., Hicks, B. M., Patrick, C. J., Carlson, S. R., Iacono, W. G., y McGue, M. (2002). Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. *Journal of Abnormal Psychology*, *111*, 411–424.
- Krueger, R. F., y Hobbs, K. A. (2020). An Overview of the DSM-5 Alternative Model of Personality Disorders. *Psychopathology*, 1–7. <https://doi:10.1159/000508538>
- Krueger, R. F., Hobbs, K. A., Conway, C. C., Dick, D. M., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Keyes, K. M., Latzman, R. D., Michelini, G., Patrick,

- C. J., Sellbom, M., Slade, T., South, S. C., Sunderland, M., Tackett, J., Waldman, I., Waszczuk, M. A., ... HiTOP Utility Workgroup. (2021). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): II. Externalizing superspectrum. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, 20(2), 171–193. <https://doi.org/10.1002/wps.20844>
- Krueger, R. F., Markon, K. E., Patrick, C. J., y Iacono, W. G. (2005). Externalizing Psychopathology in Adulthood: A Dimensional-Spectrum Conceptualization and Its Implications for DSM–V. *Journal of Abnormal Psychology*, 114(4), 537-550.
- Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., y Kramer, M. D. (2007). Linking Antisocial Behavior, Substance Use, and Personality: An Integrative Quantitative Model of the Adult Externalizing Spectrum. *Journal of Abnormal Psychology*, 116(4), 645-666.
- Krueger, R., y Markon, K. (2014). The Role of the DSM-5 Personality Trait Model in Moving Toward a Quantitative and Empirically Based Approach to Classifying Personality and Psychopathology. *Annu. Rev. Clin. Psychol*, 10, 477-501. <https://doi:10.1146/annurev-clinpsy-032813-153732>
- Krueger, R. F., Watson, D., & Widiger, T. A. (2020). The vibrant intersection of personality and psychopathology research: A special issue of the Journal of Research in Personality. *Journal of Research in Personality*, 84(103890), 103890. <https://doi.org/10.1016/j.jrp.2019.103890>
- Leyva Barajas, Y. E. (2011). Una reseña sobre la validez de constructo de pruebas referidas a criterio. *Perfiles educativos*, 33(131), 131-154.
- Lilienfeld, S. O. (1990). *Development and preliminary validation of a self-report measure of psychopathic personality.(Volumes I and II)*. University of Minnesota.
- Lilienfeld, S. y Fowler, K. (2006). The Self-Report Assessment of Psychopathy. En C. Patrick (Ed.). *Handbook of Psychopathy* (pp. 107-132). New York: The Guilford Press.
- Lilienfeld, S. y Widows, M. (2005). *Psychopathic Personality Inventory-Revised. Professional Manual*. Florida: PAR.

- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85–106.
- Linacre, J.M. (2010) When to stop removing items and persons in Rasch misfit analysis?, *Rasch Measurement Transactions*, 2010, 23:4, 1241
- Linacre, J.M. (2018). A user's guide to WINSTEPS: Rasch-model computer program. Chicago: Mesa Press.
- Linacre, J.M. (2022). A user's guide to WINSTEPS: Rasch-model computer program. Mesa Press.
- Livesley, W. J., Schroeder, M. L., Jackson, D. N., y Jang, K. L. (1994). Categorical distinctions in the study of personality disorder: Implications for classification. *Journal of Abnormal Psychology*, 103(1), 6–17. <https://doi.org/10.1037/0021-843x.103.1.6>
- Lloret, S., Ferreres, A., Hernández, A., y Tomás, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. doi: 10.6018/analesps.30.3.199361
- Lykken, D.T. (2007). Psychopathic Personality. En C. J. Patrick (Ed.), *Handbook of Psychopathy, First Edition* (1.^a ed., pp. 3–13). Guilford Publications.
- Lynam, D.R., Smith, G.T., Whiteside, S.P., Cyders, M.A., 2006. The UPPS-P: Assessing Five Personality Pathways to Impulsive Behavior. Purdue University, West Lafayette.
- Martínez-Arias, M. R., Hernández-Lloreda, M. V., y Hernández-Lloreda, M. J. (2014). *Psicometría*. Madrid: Alianza Editorial.
- Masís, A. (2008). *Ofensores sexuales extrafamiliares de personas menores de edad: Un estudio desde la perspectiva de la psicología forense en centros de atención institucional del Sistema Penitenciario Costarricense*. [Tesis de Licenciatura]. Universidad de Costa Rica, Costa Rica.
- McKinley, J., y Hathaway, S. R. (1944). The MMPI: hysteria, hypomania, and psychopathic deviate. *Journal of Applied Psychology*, 28, 153–174.

- Miguel, J. P., Silva, J. T., & Prieto, G. (2013). Career Decision Self-Efficacy Scale — Short Form: A Rasch analysis of the Portuguese version. *Journal of Vocational Behavior*, 82(2), 116–123. <https://doi.org/10.1016/j.jvb.2012.12.001>
- Ministerio de Justicia y Paz. (2018). Tablas de población del Sistema Penitenciario 2018 [Publicación web]. Recuperado de <http://mjp.go.cr/Home/Poblacion>
- Mora, C y Puercallas, P. (2005). *Homicidio y psicopatía: Tipos de violencia utilizados en el homicidio por sujetos que presentan el trastorno de personalidad psicopática*. [Tesis de Licenciatura]. Universidad de Costa Rica, Costa Rica
- Mora, C. (2013). *Simulación de síntomas psicóticos y su evaluación psicológica: Algunos elementos psicojurídicos y su relación con aspectos clínicos y forenses en el contexto costarricense*. [Tesis de posgrado]. Universidad de Costa Rica, Costa Rica.
- Morey, L. C. (2011). *Inventario de Evaluación de la Personalidad* (M. Ortiz-Tallo, P. Santamaría, V. Cardenal, M.P. Sánchez, Trad). TEA. Obra original publicada en 2007
- Mulay, A.L., Finn, M.T., Jenna, E., Gilmore, Cain, N.M., Gordon, R.M., Hopwood, C.J., & Streiner, D.L. (2019). MMPI-2 Personality Disorder Spectra Scales. *Archives of Assessment Psychology*, 9(1), 3-41.
- Mullins-Sweatt, S. N., Bornovalova, M. A., Carragher, N., Clark, L. A., Corona Espinosa, A., Jonas, K., Keyes, K. M., Lynam, D. R., Michelini, G., Miller, J. D., Min, J., Rodriguez-Seijas, C., Samuel, D. B., Tackett, J. L., y Watts, A. L. (2021). HiTOP Assessment of Externalizing Antagonism and Disinhibition. *Assessment*, 29(1), 34–45. <https://doi.org/10.1177/10731911211033900>
- Myers, N. D., Wolfe, E. W., Feltz, D. L., & Penfield, R. D. (2006). Identifying Differential Item Functioning of Rating Scale Items with the Rasch Model: An Introduction and an Application. *Measurement in Physical Education and Exercise Science*, 10, 215-240. doi:10.1207/s15327841mpee1004_1
- National Institute of Mental Health. (s.f.). Definitions of the RDoC Domains and Constructs [Publicación web]. Recuperado de <https://www.nimh.nih.gov/research-priorities/rdoc/definitions-of-the-rdoc-domains-and-constructs.shtml>

- Nichols, D. S., Kaufman, A. S., & Kaufman, N. L. (2011). *Essentials of MMPI-2 Assessment* (Second Edition). Wiley.
- Organización Mundial de la Salud. (OMS). (2001). *International Classification of functioning, disability and health*. Short version. Ginebra.
- Orozco, M. y Quesada, D. (2018). *Elaboración de un modelo de evaluación de riesgo de violencia para portación de armas de fuego*. [Tesis de licenciatura]. Universidad de Costa Rica, Costa Rica.
- Ostrosky-Solis, F., Ruiz, A., Arias, N., y Vásquez, V. (2008). Estandarización de la PCL-R en población penitenciaria mexicana. *Revista Neuropsicología, Neuropsiquiatría y Neurociencias*, 8, 49-58.
- Patrick, C. J. (Ed.). (2006). *Handbook of psychopathy*. Guilford Publications.
- Paulhus, D. L., Neumann, C. S., y Hare, R. D. (2016). *Self-Report Psychopathy Scale 4th Edition (SPR 4) Manual*. Ontario: Multi Health Systems.
- Picado, S. (2007). *Las Características Psicosociales De Los Ofensores Domésticos Homicidas De Su Pareja: Un Estudio Realizado Con Privados De Libertad En El Centro De Atención Institucional La Reforma, Centro De Atención Institucional San Rafael Y Centro De Atención Institucional Gerardo Rodríguez Echeverría*. [Tesis de Licenciatura]. Universidad de Costa Rica, Costa Rica.
- Pisoni, M. (2006). *La psicopatía como factor de riesgo para la realización de la ofensa sexual en población juvenil masculina costarricense*. [Tesis de Licenciatura]. Universidad de Costa Rica, Costa Rica.
- Plá, N. (2009). *El Modelo de Rasch aplicado a la medición multidimensional*. Paper presented at the XVI Encuentro de Economía Pública, Palacio de Congresos de Granada.
- Popham, J. (2000). *Modern educational measurement. Practical guidelines for educational leaders*. Boston: Allyn and Bacon.
- Prieto, G., & Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, (77), 65-72.
- Prieto, G., & Delgado, A.R. (2003). Análisis de un test mediante el modelo de Rasch.

- Prieto, G. & Delgado, A.R. (2007). Measuring math anxiety (in Spanish) with the Rasch Rating Scale model. *Journal of Applied Measurement*, 8, 1-12.
- Prieto, G., Delgado, A.R., Perea, M.V., García, R. y Ladera, V. (2018). Contrasting Two Verbal Fluency Scoring Systems Using the Rasch Rating Scale Model. *Actualidades en Psicología*, 32(124), 65-74.
- Psychiatry.org - Asociación Americana de Psiquiatría publica el Manual diagnóstico y estadístico de trastornos mental.* (2022). American Psychiatric Association. <https://www.psychiatry.org/newsroom/news-releases/la-asociacion-americana-de-psiquiatria-publica-el-manual-diagnostico-y-estadistico-de-trastornos-mentales-quinta-edicion-revisado-del-texto-dsm-5-tr>
- Putkonen, H., Weizmann-Henelius, G., Repo-Tiihonen, E., Lindberg, N., Saarela, T., Eronen, M., & Häkkänen-Nyholm, H. (2010). Homicide, Psychopathy, and Aging-A Nationwide Register-based Case-comparison Study of Homicide Offenders Aged 60 Years or Older*. *Journal of Forensic Sciences*, 55(6), 1552–1556. <https://doi.org/10.1111/j.1556-4029.2010.01488.x>
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of educational statistics*, 4(3), 207-230.
- Reyes, B., y Sanabria, J. (2006). *La experiencia de culpa y su relación con la atribución causal del dolor o perjuicio de los demás: Un estudio con estafadores psicopáticos y no psicopáticos.* [Tesis de Licenciatura]. Universidad de Costa Rica, Costa Rica.
- Ritchey, F. (2008). *Estadísticas para las ciencias sociales.* México, D.F: McGraw-Hill Interamericana
- Rojas, S., y Montero, E. (2012). El modelo de Rasch: una herramienta para la evaluación educativa de gran escala. *Estado de la Nación* (pp. 1-9). Costa Rica: Consejo Nacional de Rectores.
- Romero, W., y Bologna, E. (2011). Técnicas de muestreo. En E. Bologna (Ed.), *Estadística para Psicología y Educación* (1 ed., pp. 251-275). Córdoba: Editorial Brujas.

- Ruiz, M. A., Hopwood, C. J., Edens, J. F., Morey, L. C., & Cox, J. (2018). Initial development of pathological personality trait domain measures using the Personality Assessment Inventory (PAI). *Personality Disorders, 9*(6), 564–573.
- Saborío, C. (2005). Psicopatía y Violación: Un estudio de con ofensores sexuales costarricenses. *Medicina Legal de Costa Rica, 22*(1); 17-39.
- Saborío, C.; Miranda, Miranda, Mora, Retana y Rojas (2010). Inventario Multifásico de la Personalidad de Minnesota 2 (MMPI-2): Revisión y estandarización para Costa Rica. Universidad de Costa Rica.
- Saborío, C. (2011). *Extracto de: Análisis de un test para la evaluación de funciones ejecutivas (memoria de trabajo) a través del modelo de Rasch*. Universidad de Salamanca. Salamanca.
- Saborío, C. (2022). Guía de Análisis para la interpretación de resultados con el Modelo de Rasch (Manuscrito sin publicar). Universidad de Costa Rica. Salvador, B., Arce, R., Rodríguez-Díaz, F.J., y Seijo, D. (2017). Evaluación psicométrica de la psicopatía: una revisión metaanalítica. *Revista Latinoamericana de Psicología, 49*(1), 36-47. <https://dx.doi.org/10.1016/j.rlp.2015.09.015>
- Sandín, B. (2013). DSM-5: ¿Cambio de paradigma en la clasificación de los trastornos mentales?. *Revista de psicopatología y psicología clínica, 18*(3), 255-286.
- Sellbom, M., Cooke, D. J., y Shou, Y. (2019). Development and Initial Validation of the Comprehensive Assessment of Psychopathic Personality–Self-Report (CAPP-SR). *Psychology Assessment, 31*(7), 878-894.
- Sellbom, M., Liggins, C., Laurinaityté, I., y Cooke, D. J. (2021). Factor Structure of the Comprehensive Assessment of Psychopathic Personality-Self-Report (CAPP-SR) in Community and Offender Samples. *Psychological Assessment*. doi: 10.1037/pas0001029
- Sellbom, M., Lilienfeld, S. O., Fowler, K. A., y McCrary, K. L. (2018). The Self-Report Assessment of Psychopathy. En C. Patrick (Ed.). *Handbook of Psychopathy* (pp. 211-258). New York: The Guilford Press.

- Simonsen, E. (2010). The Integration of Categorical and Dimensional Approaches to Psychopathology. En T. Millon, R. Krueger y E. Simonsen. (Eds.). *Contemporary Directions in Psychopathology: scientific foundations of the DSM-V and ICD-11* (pp. 350-361). New York: The Guilford Press.
- Smith, V. y Molina, M. (2011). *La entrevista cognitiva: Guía para su aplicación en la evaluación y mejoramiento de instrumentos de papel y lápiz*. Cuadernos Metodológicos: Universidad de Costa Rica.
- Taber, K. S. (2017). *The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education*. *Research in Science Education*. doi:10.1007/s11165-016-9602-2
- Tellegen, A. (1978/1982). *Brief manual for the multidimensional personality questionnaire*. Unpublished manuscript, University of Minnesota, Minneapolis.
- Vachon, D. D., Lynam, D. R., Widiger, T. A., Miller, J. D., McCrae, R. R., & Costa, P. T. (2013). Basic Traits Predict the Prevalence of Personality Disorder Across the Life Span. *Psychological Science*, 24(5), 698–705.
<https://doi.org/10.1177/0956797612460249>
- Venables, N. C. y Patrick, C. J. (2012). Validity of the Externalizing Spectrum Inventory in a Criminal Offender Sample: Relations with Disinhibitory Psychopathology, Personality, and Psychopathic Features. *Psychology Assessment*, 24(1), 88-100.
- Widiger, T.A., Bach, B., Chmielewski, M., Clark, L.A., DeYoung, C., Hopwood, C.J., Kotov, R., Krueger, R.F., Miller, J.D., Morey, L.C., Mullins-Sweatt, S.N., Patrick, C.J., Pincus, A.L., Samuel, D.B., Sellbom, M., South, S.C., Tackett, J.L., Watson, D., Waugh, M.H., Wright, A.G.C... Thomas, K. M. (2019). Criterion A of the AMPD in HiTOP. *Journal of personality assessment*, 101(4), 345-355.
- Widiger, T. A., y Crego, C. (2019a). The Five Factor Model of personality structure: an update. *World Psychiatry*, 18(3), 271.

- Widiger, T. A., y Crego, C. (2019b). Psychopathy and DSM-5 Psychopathology. En C. J. Patrick (Ed.), *Handbook of Psychopathy, Second Edition* (2nd ed., pp. 281–296). Guilford Publications.
- Wolfradt, U., Oubaid, V., Strauber, B & Mischo, J. (1999). Thinking styles, schizotypal traits and anomalous experiences. *Personality and individual differences*. 27:821–30.
- World Health Organization. (1992). *International Statistical Classification of Diseases and Related Health Problems*. Geneva: World Health Organization.
- Wright, E. M. (2009). The measurement of psychopathy: Dimensional and taxometric approaches. *International Journal of Offender Therapy and Comparative Criminology*, 53(4), 464–481.
- Wygant, D. B., Engle, J. E., y Sellbom, M. (2020). Further examination of DSM-5 antisocial personality disorder and psychopathy: Findings from a female correctional sample. *Personality and Mental Health*. doi:10.1002/pmh.1494

Anexos

Anexo 1

Características del trastorno psicopático según Cleckley (1988).

1. Encanto superficial y buena inteligencia
2. Ausencia de delirios y otras señales de pensamiento irracional
3. Ausencia de nerviosismo o manifestaciones psiconeuróticas
4. Falta de fiabilidad
5. Falsedad e insinceridad
6. Falta de remordimiento o vergüenza
7. Comportamiento antisocial inadecuadamente motivado
8. Pobre juicio y falta de aprendizaje por experiencia
9. Egocentricidad patológica e incapacidad para amar
10. Pobreza general en relaciones afectivas mayores
11. Pérdida específica de intuición (insight)
12. Falta de respuesta en relaciones interpersonales en general
13. Comportamiento fantástico y no atractivo con la bebida y algunas veces si ésta
14. Suicidio raramente cometido
15. Vida sexual impersonal, trivial, y pobremente integrada
16. Incapacidad para seguir un plan de vida

Anexo 2

Esquema del PCL (elaboración propia) y características del trastorno psicopático según Hare (2003).

1. Encanto superficial/facilidad de palabra
2. Sentido grandioso del auto-valor
3. Necesidad de estimulación/propensión al aburrimiento
4. Mentiroso patológico
5. Tramposo/manipulador
6. Carece de remordimiento o culpa

7. Afecto superficial
8. Insensible/carencia de empatía
9. Estilo de vida parasítica
10. Pobre control conductual
11. Comportamiento sexual promiscuo
12. Problemas conductuales tempranos
13. Carencia de metas realistas a largo plazo
14. Impulsividad
15. Irresponsabilidad
16. Incapacidad para aceptar la responsabilidad de sus propias acciones
17. Muchas relaciones maritales cortas
18. Delincuencia juvenil
19. Revocación de libertad condicional
20. Versatilidad criminal

Anexo 3

Consigna para la entrevista cognitiva. Tomado de Smith y Molina (2011).

Instrucciones

Estamos realizando una serie de pruebas para ver cómo funciona un cuestionario. Para eso podrá encontrarlo en la siguiente pestaña bajo el nombre "E.C", y le voy a pedir que lo llene de forma honesta como si estuviera realizando la encuesta. En esta etapa lo que nos interesa es saber cómo está funcionando el cuestionario. Por eso le voy a pedir que conforme lo va completando piense en voz alta; y que nos explique **qué cree que se le está preguntando y qué está pensando para llegar a su respuesta**. Es decir, que diga en voz alta todo lo que se le viene a la mente conforme va completando las preguntas.

En cada pregunta, puede que le haga otras preguntas sobre la redacción de estas, las instrucciones y las opciones de respuesta. Es importante que tenga presente que nosotros queremos saber si el cuestionario funciona. No dude en decirme si algo le parece confuso y si algo se puede mejorar.

Vamos a durar aproximadamente 60 minutos en todo el proceso.

Antes de iniciar ¿tiene alguna pregunta?

¿Desea practicar? Este ejercicio no tomará más de 5 minutos, siga las instrucciones que le voy a dar a continuación.

Anexo 4

Formulario de consentimiento informado para la población forense.

UNIVERSIDAD DE COSTA RICA
VICERRECTORÍA DE INVESTIGACIÓN
COMITÉ ÉTICO CIENTÍFICO

Teléfonos:(506) 2511-4201 Telefax: (506) 2224-9367

FÓRMULA DE CONSENTIMIENTO
INFORMADO

(Para ser sujeto en una investigación)

**Creación de una prueba psicométrica de auto reporte para la evaluación
dimensional de la psicopatía en población forense y general**

Nombre de los investigadores: _____

Nombre del/la participante: _____

A. PROPÓSITO DEL PROYECTO: El presente estudio refiere a la realización de un Trabajo Final de Graduación por parte de los estudiantes de quinto año de la carrera de Psicología de la Universidad de Costa Rica, Emmanuel Artavia Picado y Jimmy Chinchilla Alvarado. Éste se lleva a cabo con la motivación de crear una prueba psicológica para evaluar características de personalidad en población general. Se espera obtener respuestas de los diferentes participantes para la validación de la prueba. La participación en el estudio requerirá que complete un cuestionario de manera virtual.

B. ¿QUÉ SE HARÁ?: Para participar en este estudio, usted podrá reunirse virtualmente con alguno de los investigadores o evaluadores. En dicha reunión, se toman datos generales sobre su persona, se aclaran dudas referentes al proyecto. Se le compartirá un enlace con el que podrá acceder al un cuestionario digital que incluye el consentimiento informado; una vez aceptado el consentimiento, se le aplicará un cuestionario cuya duración para completar se estima que ronda entre 30 minutos y 1 hora.

C. RIESGOS: La participación en este estudio puede significar cierto riesgo o molestia para usted por lo siguiente: la duración o longitud de la prueba puede percibirse como extensa y/o generar fatiga. En caso de presentarse lo anterior, se puede tomar un pequeño descanso o reprogramar la sesión para terminarla en otro momento.

D. BENEFICIOS: Como resultado de su participación en este estudio, no obtendrá ningún beneficio directo, sin embargo, es posible que los investigadores aprendan más acerca de las características de personalidad y su medición, y este conocimiento beneficie a otras personas en el futuro.

E. VOLUNTARIEDAD: La participación en el estudio es completamente voluntaria. Asimismo, puede interrumpir su participación en cualquier momento sin ninguna consecuencia.

F: CONFIDENCIALIDAD: La entrevista que se le realizará será grabada si usted autoriza al investigador a hacerlo, de lo contrario no será el caso. La información será manejada con estricta confidencialidad y anonimato. Los resultados de esta etapa se utilizarán como insumos para los siguientes pasos de la investigación. No se darán a conocer datos de los participantes específicamente.

G. INFORMACIÓN: Antes de dar su autorización para este estudio usted debe haber hablado con Emmanuel Artavia Picado o Jimmy Chinchilla Alvarado, quienes son los investigadores en este estudio y ellos deben haber contestado satisfactoriamente todas sus preguntas. Si quisiera más información más adelante, puedo obtenerla llamando a Carlos Saborío Valverde al teléfono 2511-5561, en el horario de 9:00 a.m. a 5:00 p.m.

H. Puede solicitar una copia de esta fórmula para su uso personal.

I. Su participación en este estudio es voluntaria. Tiene el derecho de negarse a participar o a discontinuar su participación en cualquier momento, sin que esta decisión afecte la calidad de la atención médica (o de otra índole) que requiere.

J. Su participación en este estudio es confidencial, los resultados podrían aparecer en una publicación científica o ser divulgados en una reunión científica, pero de una manera anónima.

K. No perderá ningún derecho legal por firmar este documento.

CONSENTIMIENTO

He leído o se me ha leído, toda la información descrita en esta fórmula, antes de firmarla. Se me ha brindado la oportunidad de hacer preguntas y éstas han sido contestadas en forma adecuada. Por lo tanto, accedo a participar como sujeto de investigación en este estudio.

__Nombre, cédula y firma del sujeto fecha

__Nombre, cédula y firma del testigo fecha

__Nombre, cédula y firma del Investigador que solicita el consentimiento fecha

NUEVA VERSIÓN FCI – APROBADO EN SESION DEL COMITÉ ÉTICO CIENTÍFICO (CEC) NO. 149 REALIZADA EL 4 DE JUNIO DE 2008.

CELM-Form.Consent-Form 06-08

Anexo 5

Formulario de consentimiento informado para la aplicación de la prueba que se pretende construir en población general.

UNIVERSIDAD DE COSTA RICA
VICERRECTORÍA DE INVESTIGACIÓN
COMITÉ ÉTICO CIENTÍFICO

Teléfonos:(506) 2511-4201 Telefax: (506) 2224-9367

FÓRMULA DE CONSENTIMIENTO
INFORMADO

(Para ser sujeto en una investigación)

**Creación de una prueba psicométrica de auto reporte para la evaluación
dimensional de la psicopatía en población forense y general**

Nombre de los investigadores: _____

Nombre del/la participante: _____

A. PROPÓSITO DEL PROYECTO: El presente estudio refiere a la realización de un Trabajo Final de Graduación por parte de los estudiantes de quinto año de la carrera de Psicología de la Universidad de Costa Rica, Emmanuel Artavia Picado y Jimmy Chinchilla Alvarado. Éste se lleva a cabo con la motivación de crear una prueba psicológica para evaluar características de personalidad en población general. Se espera obtener respuestas de los diferentes participantes para la validación de la prueba. La participación en el estudio requerirá que complete un cuestionario de manera virtual.

B. ¿QUÉ SE HARÁ?: Para participar en este estudio, usted podrá reunirse virtualmente con alguno de los investigadores o evaluadores. En dicha reunión, se toman datos generales sobre su persona, se aclaran dudas referentes al proyecto. Se le compartirá un enlace con el que podrá acceder al un cuestionario digital que incluye el consentimiento informado; una vez aceptado el consentimiento, se le aplicará un cuestionario cuya duración para completar se estima que ronda entre 30 minutos y 1 hora.

C. RIESGOS: La participación en este estudio puede significar cierto riesgo o molestia para usted por lo siguiente: la duración o longitud de la prueba puede percibirse como extensa y/o generar fatiga. En caso de presentarse lo anterior, se puede tomar un pequeño descanso o reprogramar la sesión para terminarla en otro momento.

D. BENEFICIOS: Como resultado de su participación en este estudio, no obtendrá ningún beneficio directo, sin embargo, es posible que los investigadores aprendan más acerca de las características de personalidad y su medición, y este conocimiento beneficie a otras personas en el futuro.

E. VOLUNTARIEDAD: La participación en el estudio es completamente voluntaria. Asimismo, puede interrumpir su participación en cualquier momento sin ninguna consecuencia.

F: CONFIDENCIALIDAD: La entrevista que se le realizará será grabada si usted autoriza al investigador a hacerlo, de lo contrario no será el caso. La información será manejada con estricta confidencialidad y anonimato. Los resultados de esta etapa se utilizarán como insumos para los siguientes pasos de la investigación. No se darán a conocer datos de los participantes específicamente.

G. INFORMACIÓN: Antes de dar su autorización para este estudio usted debe haber hablado con Emmanuel Artavia Picado o Jimmy Chinchilla Alvarado, quienes son los investigadores en este estudio y ellos deben haber contestado satisfactoriamente todas sus preguntas. Si quisiera más información más adelante, puedo obtenerla llamando a Carlos Saborío Valverde al teléfono 2511-5561, en el horario de 9:00 a.m. a 5:00 p.m.

H. Puede solicitar una copia de esta fórmula para su uso personal.

I. Su participación en este estudio es voluntaria. Tiene el derecho de negarse a participar o a discontinuar su participación en cualquier momento, sin que esta decisión afecte la calidad de la atención médica (o de otra índole) que requiere.

J. Su participación en este estudio es confidencial, los resultados podrían aparecer en una publicación científica o ser divulgados en una reunión científica, pero de una manera anónima.

K. No perderá ningún derecho legal por firmar este documento.

CONSENTIMIENTO

He leído o se me ha leído, toda la información descrita en esta fórmula, antes de firmarla. Se me ha brindado la oportunidad de hacer preguntas y éstas han sido contestadas en forma adecuada. Por lo tanto, accedo a participar como sujeto de investigación en este estudio.

__Nombre, cédula y firma del sujeto fecha

__Nombre, cédula y firma del testigo fecha

__Nombre, cédula y firma del Investigador que solicita el consentimiento fecha

NUEVA VERSIÓN FCI – APROBADO EN SESION DEL COMITÉ ÉTICO CIENTÍFICO (CEC) NO. 149 REALIZADA EL 4 DE JUNIO DE 2008.

CELM-Form.Consent-Form 06-08

Anexo 6

Guía de interpretación provisional de la prueba construida.

GUÍA DE INTERPRETACIÓN DE PUNTAJES DEL EDP

La Escala Dimensional de Psicopatía (EDP) es un instrumento construido en nuestro país que se encuentra en proceso de validación con población costarricense. Los constructos que mide han sido modelados a partir de los desarrollos conceptuales y empíricos de Cooke y sus colaboradores (Cooke, Hart y Logan, 2012; Cooke, Hart, Logan y Michie, 2012; Selbom, Cooke y Shou, 2019). Por lo tanto, de manera provisional se utilizan las recomendaciones de estos autores para interpretar tanto los Dominios como las distintas dimensiones (Facetas) que los componen. Se han desarrollado normas provisionales con la muestra normativa que se está terminando de completar. De esta forma, se seguirán los siguientes lineamientos para interpretar tanto los Dominios como las Facetas del EDP.

Puntajes entre 60T y 64T sugieren una elevación moderada, lo que implicaría la presencia de rasgos desadaptativos de personalidad asociados con cada uno de los constructos. Puntajes de 65T y superiores señalan una mayor severidad en la manifestación de estos rasgos.

A continuación, se reproduce la descripción proporcionada por Cooke, Hart y Logan (2007) de cada uno de los constructos de su modelo de personalidad psicopática:

I. DOMINIO APEGO

Este dominio se refiere a los vínculos emocionales que una persona tiene o es capaz de tener con los otros; es decir, el grado de afiliación, intimidad y aceptación por los otros que el individuo busca alcanzar y mantener en los intercambios interpersonales.

I a. Desapegado

- Distante

Esquivo, no amigable, inaccesible a los otros.

- Reservado

No intima, reacio en manifestar su interior, apartado o distante de los otros, desconectado de los otros; muestra que no quiere ser amistoso ni mantener relaciones con los demás.

- Frío

Falta de calidez, simpatía o afecto; poco expresivo, apático en respuesta a las necesidades, preocupaciones o incluso el disfrute de otros; no muestra emoción, compasión o bondad; no siente ni expresa simpatía o solidaridad.

I b. No adquiere compromisos

- Infiel

Hipócrita, falso, traidor, traicionero, falaz; falto de fidelidad a compromisos, deberes, creencias, individuos o ideales.

- Sin devoción hacia otros

Insolidario, carece de cualquier sentido de deber por cuidar a otros; sin entusiasmo por compromisos con los otros; no siente ni muestra amor, compromiso o lealtad hacia alguien o algo, especialmente durante mucho tiempo.

- Desleal

Innoble; indigno; muestra carencia de fe o lealtad en alguien o algo.

1.c. Sin empatía

- Sin compasión

Incomprensivo, incapaz de sentir lástima por otro, inmisericorde, sin sentimientos; no muestra sentimientos de comprensión ante el sufrimiento ajeno.

- Cruel

Indiferente al sufrimiento ajeno; malvado, malintencionado o mezquino, perverso o desagradable; poco amable; causa sufrimiento o angustia a otros deliberadamente y sin remordimientos, o es insensible al dolor y angustia de otros; ocasiona sufrimiento y malestar al otro, penoso de llevar.

- Insensible

Desalmado; sin corazón, despiadado; indiferente frente a las personas, duro; no muestra preocupación ante el daño o disgusto ajeno.

1.d Indiferente

- Desconsiderado

Egoísta, falta de consideración, no solidario, no se preocupa del otro, negligente, no piensa ni tiene respeto por otras personas o por sus sentimientos.

- Descuidado hacia los demás

No se preocupa de las consecuencias de sus actos ni de los sentimientos de los demás; no reflexiona sobre la consecuencia de sus actos; muestra una ausencia de previsión o cautela sobre las consecuencias de sus acciones en los demás.

- Negligente

Despreocupado, olvidadizo, haragán, desatento, descuidado con sus obligaciones; tiende a ignorar u olvidar sus responsabilidades hacia otros.

II. DOMINIO CONDUCTUAL

Este dominio refleja los problemas que un sujeto puede tener en la organización de actividades dirigidas hacia un fin concreto, tales como la tendencia a ser impulsivo o la búsqueda de

sensaciones. Este dominio se centra en la regulación de la conducta, incluida la capacidad del sujeto para llevar a cabo estrategias adaptativas para afrontar tareas de la vida de un modo sistemático, consistente o planificado.

II a. Inconstante

- Holgazán

Inactivo, no operativo, ocioso-desocupado; perezoso, indolente; sin metas; no desea trabajar; se dedica a actividades sin propósito serio o frívolas y pierde el tiempo.

- Indisciplinado

Rebelde, desobediente, opositor, desordenado; muestra una falta de control adecuado y de orden.

- Desidioso

Chapucero, no esmerado, no presta atención a los detalles, no le preocupa la calidad del trabajo hecho.

II b. No confiable

- Informal

Inconstante, cambiante; no se puede confiar en que actúe de la forma requerida o esperada.

- No digno de confianza

Deshonesto; marrullero; no es merecedor de confianza o de ser de confianza.

- Irresponsable

Descuidado, inmaduro, adopta medidas importantes sin la debida meditación, insensato; no tiene o no muestra cuidado alguno por las consecuencias de sus actos.

II c. Imprudente

- Precipitado

Apresurado, impulsivo, alocado, con prontos; actúa sin la debida reflexión.

- Impetuoso

Exaltado; actúa con arrebatos de energía sin considerar las consecuencias; actúa sin pensar, reaccionando a emociones o impulsos.

- Arriesgado

Busca la posibilidad de peligro (herida, perjuicio o daño); disparatado-descabellado.

II d. Intranquilo

- Excitado

Excesivamente activo, anormalmente activo; falta de capacidad para mantenerse realizando una tarea durante cierto tiempo.

- Inquieto

Moviéndose nerviosamente, incómodo, se retuerce en el asiento; se mueve o actúa de un modo repetitivo mientras mantiene básicamente la misma postura; incapacidad de relajarse en la que el alivio es buscado mediante movimientos espasmódicos.

- Vivaz

Animado, activo, despliega vigor o fortaleza, lleno de energía, lleno de impulso para actuar.

II e. Disruptivo

- Desobediente

Mala conducta, insubordinado, disconforme, desafiante; rompe reglas; rechaza o no hace lo que ordenan las leyes o quienes tienen autoridad muy habitualmente.

- Revoltoso

Desordenado, alborotador, perturbador, salvaje; difícil de controlar o disciplinar.

- Incontrolable

Inmanejable, imposible o muy difícil de tratar.

II f. Agresivo

- Amenazante

Da a entender con actos o palabras que quiere hacer algún mal a alguien, infunde miedo, asusta; manifiesta tener intención de causar daño o dolor a alguien.

- Violento

Usa o tiende a usar la agresividad física para herir a alguien o dañar algo; brutal, despiadado; poderoso.

- Intimidador

Hostigador, segrega a los otros, acosador, discriminación pernicioso; usa fuerza o poder para coaccionar a otros a través del miedo; atemoriza o maltrata a alguien más débil o en una situación de vulnerabilidad.

III. DOMINIO COGNITIVO

Este dominio se centra en procesos y acciones mentales. Se refiere a cómo el individuo centra y dirige su atención, codifica y procesa información, organiza pensamientos y hace atribuciones.

III a. Suspica

- Desconfiado

Sospecha y duda de los motivos de los otros; receloso, incrédulo; cínico.

- Cauteloso

Prudente, evita compromiso con otros; temeroso, titubeante.

- Hipervigilante

Excesivamente atento a cualquier peligro; preocupado por la necesidad de tener cuidado.

III b. Falta de capacidad de concentración

- Despistado

Cambia fácilmente de foco o desvía su atención a otra cosa.

- Desatento

No se aplica ni tiene el cuidado apropiado, no presta atención; sueña despierto.

- Descentrado

Vaguedad, poco claro, divaga, impreciso, poco específico; falta de un propósito u objetivo claro.

III c. Intolerante

- Estrecho de mente

Cerrado, sesgado; rígido o restringido a su propio punto de vista; tiene o muestra con frecuencia una visión de las cosas limitada y suele ser intolerante y con prejuicios.

- Fanático

Prejuicios irracionales; sostiene con firmeza sus opiniones, especialmente en temas como política, religión o etnias, y no quiere aceptar otros puntos de vista.

- Hipercrítico

Siempre encuentra fallos, desapueba especialmente los pequeños fallos; critica (al otro o a un hecho) demasiado y/o muy severamente.

III d. Inflexible

- Terco

Irrazonablemente obstinado, no desiste de su opinión, inamovible; persistente, emperrado, tenaz; determinado a perseverar o prevalecer irracional y obstruictivamente.

- Rígido

Estricto, severo, puntilloso; estrictamente aplicado o cumplidor, sin permisos ni excepciones; conducta, opiniones y actitudes inmodificables pese a que haya argumentos en su contra.

- Intransigente

Firme, categórico, no se doblega; no siente o no muestra deseo alguno de retractarse o llegar a un acuerdo.

III e. Ausencia de Planificación

- Sin rumbo

Vano, sin sentido, inútil, sin propósitos, sin dirección.

- Asistemático

Aleatorio, caótico, caprichoso, liado, sin método; no logra organizarse ni sigue ningún sistema.

- Desorganizado

Desordenado, incompetente.

IV. DOMINANCIA

Esta disfunción relacionada con los Trastornos de Personalidad (TP) es más evidente o llamativa en el ámbito social; la conducta interpersonal es clave para describirla. Una importante dimensión de este ámbito es la dominancia-sumisión: es decir, el grado de poder o control con el que el individuo intenta hacerse valer en las relaciones interpersonales. Las conductas interpersonales son los métodos por los cuales el individuo compromete a otros para conseguir sus propias necesidades. Un individuo que puntúa alto en este dominio entrará en frecuentes intercambios dominantes, y casi nunca tendrá una conducta sumisa, complaciente o de retirada. Su conducta es rígida e inflexible. Tienen dificultad para comprometerse en actividades de cooperación o prosociales.

IV a. Antagónico

- Hostil

Poco amigable, intimidatorio, inhospitalario, antipático, discutiendo; suele mostrar o sentir odio, enemistad o enojo.

- Desagradable

Antipático, irritante, grosero, repugnante; pendenciero, difícil, lleva la contraria, brusco o malhumorado, poco servicial; sin cortesía o siempre buscando una razón para estar en desacuerdo.

- Despectivo

Reprocha, menosprecia, condescendiente, desdeñoso, insolente; trata a los demás como si estuvieran por debajo, despreciables o merecedores de desdén; siente, expresa o demuestra un fuerte desagrado o falta de respeto por alguien o algo.

IV b. Dominante

- Arrogante

Agresivamente asertivo o presuntuoso; autoritario; engreído, altanero, egoísta, creído, superior, suficiente; siente o muestra orgullosa exaltación de sí mismo y desdén por los otros.

- Despótico

Avasallador, autoritario, pomposo; aplastante; presiona a los otros; tiende a mandar sobre los que están a su alrededor.

- Controlador

Mandón, controlador; ejerce poder o autoridad sobre otros; limita o restringe la conducta o la expresión de los otros; ejerce poder o autoridad sobre otras personas o entidades.

IV c. Engañoso

- Deshonesto

Mentiroso, hipócrita; intenta engañar, defraudar o trampear a los otros; fraudulento o insincero, falaz.

- Farsante

Induce a engaño, capaz de ser confundido por algo diferente o con otra cualidad; ilusionista; no fiable.

- Tiene dos caras

Doble juego, doble cara, traicionero, fraudulento.

IV d. Manipulador

- Tramposo

No es franco, sincero ni honesto sobre sus intenciones y motivos; injusto-desleal o poco limpio, no cumple con las reglas de juego, procedimientos o criterios usuales.

- Explotador

Injusto, abusador; usa a otros o a cosas injustamente; saca ventaja para sus propios fines.

- Calculador

Perspicaz, sagaz; intrigante; maquinador, astuto; determinado a sacar la mayor ventaja personal.

IV e. No sincero

- Superficial

Sin profundidad, rápido o ligero, aparente pero no real, sin profundidad de carácter, conocimiento o actitud, poco profundo, solo apariencia.

- Fatuo

Frívolo o pretenciosamente suave y diestro; de modales finos.

- Evasivo

No responde directamente; intenta evitar algo desagradable; furtivo, cauteloso-no suelta prenda, difícil de pillar, elude, vaguedad.

IV f. Labioso

- Con mucha labia

Fluido y locuaz, pero no sincero o no creíble, escurridizo; trivial, carece de ideas o preparación; despreocupado y relajado, con actitud informal.

- Ampuloso

Utiliza o expresa más palabras que las necesarias; lenguaje complicado o prolijo; rimbombante.

- Pretencioso

Alardea excesivamente de mérito o importancia, actuando como si fuera más valioso o especial de lo que puede justificarse o aparentando tener una imagen de sí mismo irrealmente elevada; se da una importancia o significación no justificada y a menudo se muestra sobrado o demasiado listo.

V. EMOCIONAL

Las dificultades emocionales o afectivas son componentes claves de los TP. Las emociones son sentimientos psicológicos reactivos a personas o situaciones que generalmente se acompañan de una reacción fisiológica. Las dificultades emocionales asociadas con los TP son múltiples. En el TP Psicopático las dificultades principales se relacionan con una experiencia afectiva superficial y con desregulación emocional

V a. Sin ansiedad

- Despreocupado

Nada le concierne, indolente; de modo inesperado y poco natural, no se muestra preocupado ni ansioso.

- Inalterable

No siente o no muestra miedo ni ansiedad; calmo, relajado, apacible.

- Intrépido

Carece de miedo ante el peligro o el desafío; atrevido.

V b. Sin capacidad de experimentar placer

- Pesimista

Siempre espera que suceda lo peor en cualquier situación; cree que este mundo es todo lo malo que podría ser o que las cosas tienden siempre al mal.

- Sombrío

Oscuro al punto de crear una atmósfera sombría o desalentadora; sin esperanza en ofrecer mínimas perspectivas de que las cosas mejorarán; panorama de vida indiferente.

- Poco entusiasta

No muestra entusiasmo, interés o excitación por nada; indiferente; tibio, poco impresionable.

V c. Sin profundidad emocional

- Impasible

Muestra poco o ningún sentimiento; impone la razón o el intelecto a los sentimientos; impasible, insensible, inexpresivo, sereno.

- Indiferente

No muestra preocupación ni interés en alguien o algo; sin predisposición o preferencia por una persona, grupo o cosa.

- Inexpresivo

No transmite sentimiento; vacío, facies rígida, aplanado, inexpresivo (cara de jugador de póker), soso.

V d. Sin estabilidad emocional

- Temperamental

Humor impredecible y errático; se irrita o disgusta fácilmente; tenso, cambios bruscos de carácter, volátil, explosivo.

- De humor variable

Cambios bruscos e impredecibles de estado de ánimo (alegre-malhumorado); cambios inusuales y difíciles de predecir. Muestra clara e intensamente emociones particulares, especialmente enojo o felicidad.

- Irritable

Se exaspera o enfada fácilmente; extremadamente sensible, susceptible, quisquilloso, irascible, cascarrabias.

V e. Sin remordimiento

- Sin arrepentimiento

No muestra o no siente pesar por algo mal hecho; no se excusa, sin mala conciencia, no se avergüenza.

- Sin disculpas

No muestra aflicción por acciones previas ni ofrece disculpas; no se corrige.

- Desvergonzado

Sin vergüenza ni bochorno; no se limita, restringe ni evita por un sentimiento de vergüenza o bochorno; descarado, sinvergüenza, flagrante.

VI. SÍ MISMO

El dominio 'sí mismo' se refiere a la conciencia de la propia identidad. Este dominio abarca la apreciación que el individuo tiene tanto de sus propios rasgos y esquemas como de sus habilidades, cualidades y necesidades más destacadas. La conciencia de la propia identidad también influye en los roles sociales y en cómo el individuo se relaciona con otros. El conocimiento y la conciencia de sí mismo, en interacción con las expectativas y preferencias de los otros, puede llevar a problemas para el individuo en sus relaciones. Por otra parte, un sentido de sí mismo empobrecido o conflictivo puede dificultar el proponerse metas o proyectos de futuro, por una falta de claridad y consistencia sobre los propios intereses y habilidades

VI a. Centrado en sí mismo

- Egocéntrico

Perspectiva limitada de las cosas, principalmente las relacionadas con sí mismo o los asuntos y actividades propias; ególatra, considera que es el centro del universo de cualquier experiencia.

- Egoísta

Preocupado por sus propios intereses, necesidades y deseos, ignora los de los otros; muestra que las necesidades y deseos propios deben ser vistos como más importantes que los de otras personas; desconsiderado.

- Absorto en sí mismo

Excesivamente preocupado por su propia vida e intereses; lleno de sí mismo, narcisista; exagerado sentido de su propia importancia, tiende a escribir o hablar de sí mismo todo el tiempo.

VI b. Exaltación de sí mismo

- Engreído

Tiene una alta valoración, no realista de su propia importancia o valor; pomposo; creído.

- Presuntuoso

Tiene o muestra una excesivamente alta opinión de sus cualidades o habilidades; demasiado orgulloso; presumido, altanero, estirado.

- Altivo

Actúa como si se considerara social o intelectualmente superior, especialmente cuando explica o da algo; condescendiente, altanero, snob.

VI c. Se siente único

- Se siente extraordinario

Siente que es inusual, poco común, asombroso, particular o notable; se siente poco corriente, merecedor de atención y comentarios por ser tan maravilloso, excelente, extraño e impactante.

- Se siente excepcional

Siente que tiene inteligencia o habilidad muy por encima de la media; no se ajusta a una regla o patrón general.

- Se siente especial

Se siente distinto, diferente, poco usual o superior en comparación con otros de la misma clase.

VI d. Se siente con excesivos derechos

- Demandante

Requiere mucho tiempo, atención, energía o recursos; desafiante, duro, difícil, agotador, exigente.

- Insistente

Persistente en sostener o demandar algo; demanda persistente y persuasiva de atención; implacable.

- Merecedor

Sensación de ser digno de algo por necesidad, mérito o justicia, aun cuando los demás no aprecien lo mismo; siente que merece elogios o reconocimientos por sus esfuerzos.

VI e. Se siente invulnerable

- Se siente invencible

Siente que, por su fuerza o habilidad, no puede ser vencido o derrotado; no puede ser superado.

- Se siente indestructible

Siente que es difícil o imposible que sea destruido, que es permanente o eterno.

- Se siente imbatible

Siente que es el favorito, que es demasiado bueno como para ser vencido o superado; se siente incomparable y supremo.

VI f. Autojustificación

- Minimizador

Reduce algo al mínimo grado o cantidad posible para restarle importancia.

- Negador

Se niega a aceptar la responsabilidad de sus propios actos; llegando a afirmar que no ha ocurrido o que no es el caso.

- Culpa a otros

Considera que otro es responsable de algo equivocado o desafortunado; encuentra siempre defectos en el otro.

VI g. Autoconcepto inestable

- Lábil

La visión de sí mismo o el entendimiento de su propia personalidad cambian fácilmente, a menudo con frecuencia, dramáticamente, y a la mínima provocación.

- Incompleto

Fuerte e incómodo sentimiento de malestar, de que carece de algo importante; esta carencia le aparta del prójimo, como si no perteneciera al grupo; no se siente completo ni totalmente realizado.

- Caótico

El sentido de sí mismo como persona está totalmente trastornado y fuera de control, su comprensión de sí mismo está desordenada y confusa.