DEPARTAMENTO DE PSICOLOGÍA
UNIVERSIDAD DE ALMERÍA

TESIS DOCTORAL

APROXIMACIÓN DIMENSIONAL A LA CONCEPTUALIZACIÓN Y MEDICIÓN DEL FENOTIPO AMPLIADO DEL AUTISMO

DIMENSIONAL APPROACH TO THE CONCEPTUALIZATION AND MEASUREMENT OF THE BROAD AUTISM PHENOTYPE

MARTA GODOY GIMÉNEZ

TESIS DOCTORAL DIRIGIDA POR:

MARIA DE LOS ÁNGELES FERNÁNDEZ ESTÉVEZ, FERNANDO CAÑADAS PÉREZ Y PABLO SAYANS JIMÉNEZ

ESCUELA INTERNACIONAL DE DOCTORADO

DOCTORADO EN SALUD, PSICOLOGÍA Y PSIQUIATRÍA

ALMERÍA A ENERO DE 2022

"A los mayores les gustan las cifras. Cuando se les habla de un nuevo amigo, jamás preguntan sobre lo esencial del mismo. Nunca se les ocurre preguntar: "¿Qué tono tiene su voz? ¿Qué juegos prefiere? ¿Le gusta coleccionar mariposas?" Pero en cambio preguntan: "¿Qué edad tiene? ¿Cuántos hermanos? ¿Cuánto pesa? ¿Cuánto gana su padre?" Solamente con estos detalles creen conocerle."

El Principito, Antoine de Saint-Exupéry

"I began to wonder about the impact of a variety of methodological factors that might have influenced the results: … I wondered about the way that the test had been translated from English to the other languages. International studies of educational achievement can be invaluable to policy makers and educators but not if methodological factors undermine the validity of the results."

Translating and Adapting Tests for Cross-Cultural Assessments, Ronald K. Hambleton

"Determining whether construct equivalence exists between two cultures involves primarily judgmental strategies. A researcher must begin by using his or her common sense to answer such questions as, Is it sensible to compare these two cultures on this construct? Does the construct that is being measured have similar meaning in all cultures being compared?"

Translating and Adapting Tests for Cross-Cultural Assessments, Ronald K. Hambleton

La presente tesis doctoral ha sido tutorizada por la Dr. María de los Ángeles Fernández Estévez, catedrática del Área de Psicología Básica de la Universidad de Almería y codirigida por el Dr. Fernando Cañadas Pérez, profesor titular del Área de Psicobiología y por el Dr. Pablo Sayans-Jiménez, profesor del Área de Metodología de las Ciencias del Comportamiento de la misma universidad.

Su defensa tendrá lugar en el Departamento de Psicología de la Universidad de Almería.

Reconocimientos

Los estudios incluidos en la presente tesis doctoral han sido parcialmente financiados por dos proyectos nacionales concedidos a la tutora académica de la doctoranda (las referencias de dichos proyectos son PSI2015-65248-P y PID2019-110066GB-I00/MCIN/AEI/ 10.13039/501100011033). Asimismo, la autora de la tesis ha contado con una ayuda predoctoral para la Formación del Profesorado Universitario (FPU) al cargo del Ministerio Español de Educación, Cultura y Deporte (referencia de la ayuda FPU15/01562). Las fuentes de financiación arriba citadas no han intervenido en el diseño y/o programación de los estudios, ni en los procesos de recogida, análisis e interpretación de los datos, así como tampoco en la preparación de los correspondientes informes o manuscritos científicos ni en la decisión de enviarlos a revistas del ámbito científico para su posible publicación. La autora declara no tener otros conflictos de intereses.

Acknowledgments

The studies included in the present dissertation have been partially funded by two national projects granted to the academic supervisor of the author (references are PSI2015-65248-P y PID2019-110066GB-I00/MCIN/AEI/ 10.13039/501100011033) by the Spanish Ministry of Economy and Competitiveness, co-funded by European ERDF funds. The author of the dissertation was supported by a predoctoral grant of the Formation of the University Staff by the Ministry of Education, Culture, and Sport (grant reference FPU15/01562. The funding sources abovementioned have not taken part in the design or planification of the studies, nor in the data collection, analysis and interpretation or in the preparation of final reports and scientific manuscripts and in the decisions of submission for their consideration for publication in scientific journals. The author declares no conflicts of interest to disclosure.

En la presente tesis doctoral se opta por una mención internacional a doctora. Para tal efecto, las siguientes páginas del presente documento han sido escritas únicamente en lengua inglesa a salvedad de una introducción de la misma en español.

DIMENSIONAL APPROACH TO THE CONCEPTUALIZATION AND MESUAREMENT OF THE BROAD AUTISM PHENOTYPE

Contents

List of Tables

List of Tables

List of Figures

List of Figures

List of abbreviations

ADI—R	Autism Diagnostic Interview
ADOS—2	Autism Diagnostic Observation Schedule
AERA	American Educational Research Association
AQ	Autism Quotient
ASD	Autism Spectrum Disorder
APA	American Psychiatric Association
BAP	Broad Autism Phenotype
BAP-IT	Broad Autism Phenotype—International Test
BAPQ	Broad Autism Phenotype Questionnaire
BPASS	Broader Phenotype Autism Symptom Scale
CAST	Childhood Asperger Syndrome Test
CCC-2	Children's Communication Checklist—2
CFA	Confirmatory Factor Analysis
CID	International Statistical Classification of Diseases and Related Health Problems
CVI	Content Validity Index
DSM	Diagnostic and Statistical Manual of Mental Disorders
DFT	Differential Test Functioning
DIF	Differential Item Functioning
EFA	Exploratory Factor Analysis
E/CFA	Exploratory Factor Analysis in the Confirmatory Factor Analysis Framework
e.g.	Exampli gratia (for example)
-EN	English version of a test
EN	English Group of Participants
ENA	Empirical Network Analysis
ES-	Spanish Group of Participants
FAA	Fenotipo Ampliado del Autismo
i.e.	Id est (that is)
IQ	Intelligence Quotient
IRT	Item Response Theory

IRI	Interpersonal Reactivity Index
ITC	International Test Commission
MIs	Modification Indices
MLR	Maximum Likelihood Estimation with Robust Standard Errors
NCME	National Council on Measurement in Education
OCD	Obsessive-Compulsive Disorder
OCI-R	Obsessive-Compulsive Inventory-Revised
PAQ	Pragmatic Awareness Questionnaire
PRS	Pragmatic Rating Scale
RA	Stereotyped or Repetitive Motor Movements, Use of Objects, or Speech
RB	Insistence on Sameness, Inflexible Adherence to Routines, or Ritualized Patterns or Verbal and Nonverbal Behaviour
RC	Highly Restricted, Fixated Interests that are Abnormal in Intensity or Focus
RCS	Resistance to Change Scale
RD	Hyper- or Hyporeactivity to Sensory Input or Unusual Interests in Sensory Aspects of the Environment
RIRE-BAP	Restricted and Repetitive Patterns of Behaviours Dimension of the BAP-IT
RRBs	Restricted, Repetitive Behaviour
SA	Deficits in Social-Emotional Reciprocity (ASD subdomain)
SATQ	Subthreshold Autism Trait Questionnaire
SB	Deficits in Nonverbal Communicative Behaviours Used for Social Interaction (ASD subdomain)
SC	Deficits in developing, Maintaining, and Understanding Relationships (ASD subdomain)
SCI	Social Communication and Interaction
SCDC	Social and Communication Disorders Checklist
SCQ	Social Communication Questionnaire
SOCIAL-BAP	Social Interaction and Social Communication Deficits Dimension of the BAP-IT
-SP	Spanish version of a test
SPIN	Social Phobia Inventory

SRS—II	Social Responsiveness Scale—Second Edition
TEA	Trastorno del Espectro Autista
WHO	World Health Organization

Introducción

En la presente tesis doctoral ha convergido el conocimiento de tres áreas fundamentales de la psicología con el objetivo de demostrar que la investigación transversal no es solo posible sino también altamente enriquecedora para nuestra disciplina. La psicología cognitiva representa la línea base de esta tesis puesto que mi propósito inicial era explorar experimentalmente los procesos cognitivos involucrados en el fenotipo ampliado del autismo (FAA)—especialmente aquellos relacionados con la cognición social. De ahí, el interés por la psicobiología se alcanza de manera natural ya que también me interesaron los mecanismos neurobiológicos y los circuitos neuronales que subyacen a cada proceso cognitivo y comportamiento prototípico del FAA y cómo estos se relacionaban unos con otros. Por último, pero no menos importante, esta tesis doctoral debe gran parte de su esencia a la psicometría. Al respecto de esto, se ha defendido un enfoque integrador donde las teorías sobre los fenómenos psicológicos y psicométricos han trabajado de la mano con el objetivo de mejorar las definiciones operativa y relacional del FAA y su proceso de medición. El desarrollo de la investigación presentada en las siguientes páginas me ha llevado a asignar a la psicometría la primera posición. Su papel en mi tesis ha sido notable, puesto que se ha encargado de reconectar cómo las personas exhiben ciertos comportamientos del FAA con un modelo teórico y psicométrico capaz de representar una operacionalización del constructo en la población general española e inglesa y una herramienta de medición.

Sin embargo, para comprender el origen de esta tesis—es decir, su particularidad— y hacerla más accesible al lector, ha sido necesario entrar en el marco teórico donde se inserta: la existencia de expresiones consideradas prototípicamente autistas en poblaciones no clínicas. A la presencia de esos comportamientos agrupados se les ha designado como el FAA, fenómeno que se describió en un primer momento en familiares de primer grado de niños diagnosticados con trastorno del espectro autista (TEA) y cuyo estudio también puede ayudar a comprender las rutas genéticas del trastorno. El FAA fue documentado por primera vez hace unos setenta años (Kanner, 1943) y formalmente delineado durante los años noventa (ver Piven et al., 1997a, b). Como expresión fenotípica del autismo, la definición inicial del FAA se equiparó con el trastorno autista (TA) en la cuarta edición del Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-IV; Asociación Americana de Psicología [APA], 1994): interacción social deteriorada, daño cualitativo en la comunicación y patrones de comportamiento, intereses y actividades restringidos, repetitivos y estereotipados. Sin embargo, mientras que el TA ha progresado hacia un enfoque dimensional, el TEA (DSM-V; APA, 2013), el FAA no ha seguido el mismo camino. Por consiguiente, los hallazgos de los estudios realizados con poblaciones que expresan comportamientos considerados típicos del TEA y del FAA[1] han sido difícilmente comparables (a pesar de esto, se han realizado conexiones directas entre ellos en la literatura). Además, las comparaciones de estudios realizados solo en poblaciones en el nivel de expresividad del FAA (es decir, poblaciones que presentan, al menos, un conjunto de indicadores con un nivel de severidad asociado al FAA) podrían ser problemáticas ya que los test disponibles para evaluar el FAA no comparten la operacionalización actual del fenotipo.

La falta de ajuste entre la operacionalización del FAA y el enfoque dimensional del TEA tampoco ha impedido el crecimiento de este prolífico campo de investigación. En este sentido, el FAA ha sido ampliamente explorado en todo el mundo y han surgido varias herramientas de medición para captar su esencia. Pero, de nuevo, esas herramientas no solo no comparten la estructura dimensional más reciente del TEA, sino que ni tan siquiera comparten una estructura común para el FAA, ya que, como se revisará más adelante, cada test ha conceptualizado y evaluado el fenotipo de forma diferente.

Con todo, puedo concluir y adelantar que esta tesis apunta en primer lugar a la falta de acuerdo teórico entre el TEA y su expresión fenotípica, el FAA. En segundo lugar, aborda, asimismo, la falta de acuerdo en las operacionalizaciones sobre el FAA en las herramientas más comúnmente aplicadas para evaluarlo. En tercer lugar, muestra cómo esta discrepancia también ha propiciado la inestabilidad de las estructuras internas de dichos test conforme a distintas aplicaciones y adaptaciones (ya sean lingüísticas o de formato) de los mismos. En cuarto y último lugar, se dan las

[1] Una población TEA es aquella que expresa comportamientos autistas a nivel de severidad propio de un diagnóstico formal de TEA. Una población FAA engloba a aquellas personas de la población comunitaria que han puntuado alto en un test que mida FAA (i.e., expresan comportamientos autistas a nivel de severidad subclínico).

razones por las que la reconciliación entre los constructos del TEA y del FAA implicará la readaptación de los test disponibles para su medición o el desarrollo de nuevas herramientas en su defecto.

Por último, no podría pasar por alto el hecho de que la presente tesis doctoral ha sido desarrollada parcialmente en español y con poblaciones españolas. En este sentido, los test disponibles para la evaluación del FAA han sido extensamente aplicados y adaptados a diferentes idiomas, pero, desafortunadamente, hasta el momento en que se inició esta investigación no existía ninguna versión en español de ninguna de las herramientas de medida del FAA. Por esa razón, y dado que las pruebas disponibles se han aplicado ampliamente, la primera tarea ha sido adaptar una de dichas pruebas al español.

Para tal efecto, se ha escogido el Broad Autism Phenotype Questionnaire (BAPQ; Hurley et al., 2007) por ser la única prueba diseñada en su origen para evaluar el fenotipo y replicar la operacionalización más aceptada del FAA en tres dominios. Las propiedades psicométricas de su versión en español también han sido analizadas a través de diferentes enfoques. En particular, se han proporcionado evidencias adversas basadas en la estructura interna del BAPQ, tanto en su versión en español como en la original, que han sacado a la luz las discrepancias entre los constructos del TEA y del FAA y ha supuesto un apoyo a la redefinición del fenotipo.

Como consecuencia directa, mi investigación se ha enfocado en proponer una nueva conceptualización del FAA, alineada con la del TEA, pasando a agrupar y ordenar en esta nueva estructura del FAA los ítems de los test más comúnmente aplicados. Sin embargo, tras este trabajo, se ha observado que sigue existiendo cierta infrarrepresentación de algunos de los comportamientos más centrales del FAA. Estos hallazgos, y el respaldo de los expertos en el área, han llevado a desarrollar un nuevo test para la evaluación de la operacionalización entorno a dos dimensiones del FAA con ítems generados por la presente autora de la tesis. Dichos ítems son susceptibles de ser graduados en función de la severidad de su expresión según el enfoque dimensional del autismo, el cual ha primado en esta tesis doctoral. Además, los ítems han sido desarrollados paralelamente en dos idiomas, español e inglés, y han demostrado ser métricamente invariantes entre ambos idiomas.

Planificación del trabajo

La presente tesis doctoral se organiza en tres bloques de contenido. El primer bloque corresponde al marco teórico y se divide en dos grandes capítulos, el primero de los cuales, "Conceptualización y evaluación del fenotipo ampliado del autismo como las expresiones fenotípicas leves del autismo" comienza abordando el TEA (primera parte), su definición diagnóstica, su etiología y las aplicaciones derivadas de su estudio. Después, se pasa a hablar de su expresión fenotípica, el FAA (segunda parte). Tras caracterizar los comportamientos y procesos cognitivos más centrales del FAA, se revisa su prevalencia en familiares de primer orden de personas diagnosticadas con TEA y a través de la población general para, finalmente, hablar sobre las ventajas de su estudio. En la tercera parte de este capítulo se repasa la historia de la medición y evaluación del FAA y se exponen los problemas psicométricos y de medición más relevantes derivados de una incorrecta operacionalización del constructo. Finalmente, en la cuarta parte se sugiere el camino que debería seguir la investigación tanto en lo referente a la operacionalización actualizada del FAA como en su medición para resolver los problemas existentes y salvar la brecha entre ambos constructos. A este capítulo le sigue el segundo "Objetivos generales y específicos, enfoques de investigación y justificación de la presente tesis" que corresponde con el establecimiento de los objetivos de la tesis y del plan de investigación que se va a seguir.

Llegados al segundo bloque nos adentramos en los estudios empíricos que comprenden los cuatro estudios realizados durante mi labor como doctoranda. Los estudios están interrelacionados de manera que las conclusiones de uno han generado los motivos del siguiente. Los nombres de los capítulos de este bloque corresponden a los títulos de los diferentes estudios: el tercer capítulo "Propiedades psicométricas de la versión española del Broad Autism Phenotype Questionnaire: fortalezas, debilidades y futuras mejoras" (traducción literal de su título en inglés; publicado) presenta la adaptación al español de la prueba original del FAA, el BAPQ, y analiza sus propiedades psicométricas aportando evidencias basadas en la estructura interna de las puntuaciones de la versión en español del BAPQ y en su relación con otras variables. Las evidencias adversas basadas en la estructura interna de la versión española del BAPQ (BAPQ-SP) indicaron problemas de unidimensionalidad en una de las subescalas del test, Pragmática del Lenguaje. En el cuarto capítulo "Aplicación de un modelo de Rasch al BAPQ: análisis de la severidad de los ítems y funcionamiento diferencial del test entre las versiones inglesa y española" (en fase de revisión en una revista), considerando el enfoque dimensional del autismo propuesto en el DSM-V propongo que éste pueda extenderse a su fenotipo, el FAA y, a través de los modelos de la teoría de respuesta al ítem (TRI), analizo las propiedades psicométricas del BAPQ (tanto del BAPQ-SP como del original). Además, se explora la escalabilidad de los ítems de las dos dimensiones que no presentan problemas de dimensionalidad, Reticencia Social y Rigidez Mental y Comportamental, y se intentan construir dos continuos de severidad gradual (uno para cada una de las dimensiones). Finalmente, se indaga en el funcionamiento diferencial de las escalas del BAPQ y de sus ítems entre versiones. Los resultados de este capítulo también sugieren problemas de unidimensionalidad en la subescala de Pragmática del Lenguaje tanto en el BAPQ como en el BAPQ-SP y, por ende, no es posible continuar con el análisis de esta subescala a partir de este punto. Otros hallazgos en las dos subescalas restantes (Reticencia Social y Rigidez Mental y Comportamental) sugieren que sus ítems son altamente severos y que se requeriría incluir en las mismas ítems más leves. Asimismo, se observan diferencias en las jerarquías de severidad y en la intensidad de los ítems entre las versiones del test y se discute el papel de la influencia cultural en la expresión modulada del autismo.

Gracias a los hallazgos de estos dos primeros estudios también se hace visible la discrepancia existente entre las conceptualizaciones del TEA y del FAA y que dichas disimilitudes deberían abordarse con urgencia. Por consiguiente, en el quinto capítulo "¿Es posible evaluar el fenotipo ampliado del autismo según la definición del DSM-V del trastorno del espectro autista utilizando las herramientas de medición disponibles?" (publicado) intento armonizar las definiciones del TEA y del FAA para, a continuación, agrupar los ítems de los test más ampliamente aplicados en la evaluación del fenotipo y reestructurarlos en la nueva definición operacional que propongo. Con esto, pretendo proporcionar ítems que se ajusten a la estructura entorno a dos dimensiones del FAA. Además, todos los ítems son evaluados por un extenso grupo de profesionales y académicos expertos en autismo. Dichos expertos se centran en comprobar la relevancia de cada contenido de cada ítem para medir el FAA y la representación final del constructo. Sin embargo, si bien los ítems finales parecen ser relevantes para la medición del fenotipo, no resultan suficientes y, por lo tanto, sigue existiendo cierta infrarrepresentación de algunos de los comportamientos centrales del FAA.

Finalmente, se da un paso más y se desarrolla un nuevo test para medir el FAA con el que se pretenden cubrir esas lagunas de contenido. En el sexto capítulo, "El Broad Autism Phenotype—International Test (BAP-IT): un test para la evaluación del fenotipo ampliado del autismo basado en sus dos dimensiones" (en preparación) presento finalmente al BAP-IT, un test que evalúa la operacionalización de dos dimensiones del FAA con ítems que expresan diferentes grados de severidad según el enfoque dimensional del autismo y que está disponible en dos idiomas, español e inglés. Un grupo de expertos en autismo examina la relevancia de los ítems creados y la representación del FAA antes de dar forma al test final y aportar evidencias basadas en su estructura interna y en su relación con otras variables.

Una vez expuestos los estudios empíricos, el tercer bloque de esta tesis "Discusión de los resultados generales, conclusiones y futuras direcciones" expone una profunda discusión sobre los resultados de todos los estudios y extrae conclusiones generales sobre todos ellos. También se proponen varias posibles líneas para futuros estudios.

Introduction

The present dissertation has converged the knowledge of three fundamental areas of psychology, aiming to demonstrate that transversal research could be not only possible but also highly nurturing of our discipline. Cognitive psychology represents the baseline of this work because my initial purpose was to explore experimentally the specific cognitive processes that might be involved in the broad autism phenotype (BAP)—especially those related to social cognition. Since I was also interested in the neurobiological mechanisms and neural circuits underlying each cognitive process and BAP specific behaviour and how the ones are related to the others, an interest in psychobiology naturally arises. Last but not least, this dissertation lies its essence in psychometrics. Regarding this, an integrative approach has been defended where theories about psychological phenomena and psychometrics have worked hand in hand for improving the operational and relational definitions of the BAP and its measurement process. The justification of the research presented in the following pages has driven me to allocate psychometrics at the forefront of this thesis. Its role in my dissertation has been to reconnect how people show certain BAP behaviours with a theoretical and psychometric model capable of representing an operationalization of the construct in the Spanish and English general populations, and develop a new measurement tool.

Nevertheless, to understand the origin of the present work—id est, its spirit—and make it more accessible for the reader, it has been necessary to go into the theoretical framework where it is inserted: the existence of milder autistic expressions in non-clinical populations. The presence of those behaviours has been identified as the BAP, which was primarily observed in first-degree relatives of children diagnosed with autism spectrum disorder (ASD) and which study may also help to understand the genetic paths of the disorder. The BAP was firstly reported about seventy years ago (Kanner, 1943) and formally delineating during the nineties (see Piven et al., 1997a, b). As the autistic phenotypical expression, the initial definition of the BAP was equivalated with the autism disorder in the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association [APA], 1994): impaired social interaction, qualitative damage in communication, and restricted, repetitive, and stereotyped patterns of behaviour, interests, and activities. However, while autism disorder research has progressed towards a dimensional approach, the ASD (DSM-V; APA, 2013) the BAP-focused research has not. Thus, findings from studies carried out with ASD and BAP populations[2] have been hardly comparable (despite this, direct connections between them have been theorized in the literature). Furthermore, comparisons of studies conducted only in populations at BAP levels of expression (i.e., populations presenting, at least, a set of indicators associated with the BAP) could be problematic since the available BAP tests do not match with the current BAP operationalization.

The lack of fit between the BAP operationalization and the ASD dimensional approach has not impeded the growth of this prolific research field. In this regard, the BAP has been extensively explored worldwide and several measurement tools have emerged to grasp its essence. But again, those measurement tools do not share either the newest ASD structure or even a common BAP one since each test conceptualises and assesses the phenotype differently.

Consequently, I can conclude and advance that this dissertation targets, in the first place, the lack of theoretical agreement between the ASD and its phenotype, the BAP. In the second place, it also addresses the existence of a lack of agreement regarding how the BAP is operationalised in the most common tests applied to assess it. In the third place, it shows how this lack of agreement has also aroused the instability of the internal structures of the same tests across applications and adaptations (either to other languages or test formats). Finally, I provide the reasons why the reconciliation of both the ASD and BAP is going to imply the readaptation of the available tests or the development of new ones.

Finally, I cannot disregard the fact that the present dissertation has been partially conducted in Spain and targeting Spanish populations. In this sense, available tests to assess the BAP have been extensively applied and adapted to different languages but then again, by the time the present dissertation started, there was not a version of any test adapted to the Spanish population. Ideally, as available tests have been broadly applied in the international context, the first task within this dissertation has been to adapt one of them to the Spanish language.

[2] ASD population refers to a population composed by people diagnosed with ASD. BAP population refers to community people scoring high in a BAP test.

The chosen test has been the Broad Autism Phenotype Questionnaire (BAPQ; Hurley et al., 2007) since it is the test originally designed to assess the phenotype and it replicates the most accepted BAP structure across three domains. The psychometric properties of its Spanish version have also been studied through different approaches. Specifically, adverse evidence based on the BAPQ internal structure has been reported in either the Spanish or the original version, uncovering the discrepancies between the ASD and BAP constructs and supporting the redefinition of the latter.

As a direct consequence, my research has focused on proposing a new BAP operationalization, aligned with the ASD one, and the items of the already existing BAP tests have been regrouped and organised into the new BAP structure. At that point, and after arduous work, it still existed some underrepresentation of certain BAP core behaviours. These findings and the support of autism experts have led to the development of a new test for the assessment of the two-domains structure of the BAP, with items generated by the author of the present dissertation. These items are susceptible to be scaled in different degrees of severity according to the dimensional approach of autism defended in this dissertation. Furthermore, the items have been developed parallelly in two different languages and have been demonstrated to be metrically invariant between countries.

Work Planning

The present dissertation is organized into three blocks. The first one corresponds to the theoretical framework, and it is divided into two chapters where the first one "Conceptualization and assessment of the broad autism phenotype as the mild autistic phenotypic expressions" begins by addressing the ASD (first part), its diagnostical definition, its aetiology, and the applications of studying it. Once I have stated the ASD, I discuss its phenotypical expression, the BAP (second part). After characterising the BAP behaviourally and neuropsychologically, its prevalence in first-degree relatives and across the general population is discussed to finally address the advantages of studying the BAP. The third part of this chapter reviews the measurement history of the BAP and states the psychometrical and measurement problems derived from an incorrect operationalisation of the construct. Finally, the fourth part suggests the natural path that future research should follow both in the updated operationalization of the BAP and in its assessment to solve the problems and bridge the gap between both constructs. Continuing with the second chapter "General and specific objectives, research approaches, and justifications of the present dissertation", it corresponds to the establishment of the objectives of the dissertation and the research plan.

The second block deepens into the empirical studies, and it includes the four studies conducted during my PhD apprenticeship. Studies are interrelated in a way that the findings of one have led to the goals of the subsequent study. The names of the chapters in this block correspond to the titles of the different studies: the third chapter "Psychometric Properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements" (published work) presents the Spanish adaptation of the original test for assessing the BAP, the BAPQ, and the analysis of its psychometric properties, providing evidence based on the BAPQ scores' internal structure and on their relationship with other variables. Adverse evidence based on the Spanish BAPQ (BAPQ-SP) internal structure uncovered unidimensionality problems in one of the BAPQ subscales, Pragmatic Language. In the fourth chapter "Applying a Rasch Model to the BAPQ: Item Severity Analysis and Test Differential Functioning of the English and Spanish Versions" (work under review in a scientific journal) I consider the dimensional approach to autism proposed in the DSM-V and suggest that it can be extended to its phenotype, the BAP, and through item response theory (IRT) models, I analyse the psychometric properties of the BAPQ (both BAPQ-SP and the original BAPQ). Additionally, I explore the scalability of the items of the two domains that did not present dimensionality problems along continuums of gradual severity. As a further step, I inquire as well into the test and item differential functioning between versions. Findings from this chapter also suggested unidimensionality problems of the Pragmatic Language subscale either in the original BAPQ or the BAPQ-SP and, thus, analysis of this subscale ended at this point. Further results in the remaining two subscales (Aloof and Rigid) proposed that the items of the BAPQ were highly severe and milder items were required. Differences in the severity hierarchies and the severity of the items between versions were also observed, and the cultural modulation in the expression of autism was discussed.

Thanks to the conclusions of these two studies, a discrepancy in the conceptualizations of the ASD and the BAP is revealed and I realise that it should be urgently addressed. Consequently, in the fifth chapter “Is it possible to assess the Broad Autism Phenotype according to the current DSM-V definition of autism spectrum disorder using the available measurement tools?” (published work) I try to reharmonise the ASD and BAP definitions and, then, cluster the items of the most commonly applied tests to assess the BAP and reshape them in the new BAP operationalization and two-domains structure. With this, I attempt to provide items with content that fits the two-domais structure of the BAP. Besides, they are evaluated by a broad group of professional and academic experts in autism regarding their relevance and the BAP representativeness. Experts suggested that the remaining items were relevant for assessing the BAP but resulted insufficiently and, thus, it still existed some underrepresentation of certain BAP core behaviours.

Finally, I go a step further developing a new BAP test where those gaps have been eventually filled. In the sixth chapter, “The Broad Autism Phenotype - International Test (BAP-IT): a two-domain based test for the assessment of the broad autism phenotype” (work under review by the authors), I present the BAP-IT, a test that assesses the two-domains operationalization of the BAP with items expressing different degrees of severity according to the dimensional approach of autism and in two different languages, Spanish and English. A group of experts examine the relevance of the created items and the representation of the BAP before conforming to the final test and providing evidence-based on the BAP-IT internal structure and its relation with other variables.

Once the empirical studies have been presented, the third block “Discussion of the overall outcomes, conclusions, and future directions” exposes a general discussion of the results of the present dissertation and draws certain conclusions. It also states the research limitations and proposes several directions for future studies.

BLOCK ONE:

THE THEORETICAL FRAMEWORK

CHAPTER ONE. CONCEPTUALIZATION AND ASSESSMENT OF THE BROAD AUTISM PHENOTYPE AS THE MILD AUTISTIC PHENOTYPIC EXPRESSIONS

First Part. The Hypothesis of the Spectrum of Autism

> Autism is not only a disorder or a label that segregates the population into two categories of clinical and unclinical individuals. Autism is a wide group of behaviours and personality expressions that spreads along all the population with different intensities. When those expressions hinder the life of a person, they reach a clinical threshold and conduct to the diagnosis of autism spectrum disorder (ASD). Below that threshold, they are also clustered and known as the broad autism phenotype (BAP). The first part of this chapter addresses the clinical expressions of autism.

1.1. Novelties in the Updated ASD Definition: Two Core Domains and a Continuum of Gradual Severity

The Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder defined in the DSM–V (APA, 2013) by the presence of social communication and social relation deficits (criterion A; see Table 1) alongside a pattern of restricted repetitive behaviours and interests (criterion B; see Table 1). The current ASD definition proposed in the updating of both the DSM–V; APA, 2013) and supported by the forthcoming International Statistical Classification of Diseases and Related Health Problems (ICD-11; World Health Organization [WHO], 2019, 2022) carries some substantial changes (Constantino & Charman, 2016) in its conceptualization such as the collapsing of symptom criteria from three domains to two. This was made partly based on (i) evidence that the socio-communicative deficits commonly presented in ASD (e.g., impairment in reciprocal social interaction and impairment in social aspects of communication) are closely inter-related as well as their severity is highly correlated within clinically affected populations (Gotham et al., 2012), and also in the general one (Constantino, 2011), and (ii) evidence pointing towards connections between lack of flexibility in communication (e.g., delayed echolalia, repetitive speech and questions, and verbal rituals) and insistence on sameness (criterion B; Lam et al., 2008; Szatmari et al., 2006).

Table 1
DSM-V Diagnostic Criteria for Autism Spectrum Disorder 299.00 (F84.0)

A)	**Persistent deficits in social communication and social interaction across multiple contexts**, as manifested by the following, currently or by history (examples are illustrative, not exhaustive): **1 Deficit in social-emotional reciprocity**, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions. **2 Deficits in non-verbal communicative behaviours used for social interaction**, ranging, for example, from poorly integrated verbal and non-verbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and non-verbal communication. **3 Deficits in developing, maintaining, and understanding relationships**, ranging, for example, from difficulties adjusting behaviour to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to the absence of interest in peers.
B)	**Restricted, repetitive patterns of behaviour, interests, or activities**, as manifested by at least two of the following, currently or by history (examples are illustrative, not exhaustive): **1 Stereotyped or repetitive motor movements, use of objects, or speech** (e.g., simple motor stereotypies, lining up toys or flipping objects, echolalia, idiosyncratic phrases). **2 Insistence on sameness, inflexible adherence to routines, or ritualised patterns of verbal or non-verbal behaviour** (e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take the same route or eat the same food every day). **3 Highly restricted, fixated interests that are abnormal in intensity or focus** (e.g., strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interests). **4 Hyper- or hyporeactivity to sensory input or unusual interest in sensory aspects of the environment** (e.g., apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement).
C)	Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities or may be masked by learned strategies in later life).
D)	Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning.
E)	These disturbances are not better explained by intellectual disability (intellectual developmental disorder) or global developmental delay. Intellectual disability and autism spectrum disorder frequently co-occur; to make comorbid diagnoses of autism spectrum disorder and intellectual disability, social communication should be below that expected for the general developmental level.
	Specify if: • With or without accompanying intellectual impairment • With or without accompanying language impairment • Associated with a known medical or genetic condition or environmental factor • Associated with another neurodevelopmental, mental, or behavioural disorder • With catatonia (refer to the criteria for catatonia associated with another mental disorder)

Note. Individuals with a well-established DSM-IV diagnosis of autistic disorder, Asperger's disorder, or pervasive developmental disorder not otherwise specified should be given the diagnosis of autism spectrum disorder. Individuals who have marked deficits in social communication, but whose symptoms do not otherwise meet the criteria for autism spectrum disorder, should be evaluated for social (pragmatic) communication disorder

Adapted from the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-V; pp. 50-51), American Psychiatric Publishing, a Division of American Psychiatric Association, Arlington, V.S.

Autism is known as a “spectrum” disorder because there is a wide variation in severity of symptoms people experience, their intellectual level, and functional disability (Geschwind, 2009). Accordingly, the updated ASD definition also includes new severity specifiers (see Table 2; APA, 2013; Constantino & Charman, 2016) which aim to categorise the impact of the symptoms on the adaptative functioning of the individual. These specifiers represent the effect of symptoms in each domain on three categories of adaptative functioning, each of which incorporates descriptive scoring anchors that recommend the level of external help that a person with a diagnosis would require. Yet, its biggest relevance concerns the suggestions that even if symptoms are presented and noticeable, impairment could be relatively mild, and vice versa (very few but severe symptoms could hinder the life areas of the individual). Thus, ASD-specific types of expressions and their severities could represent orthogonal axes of diagnosis, highlighting the importance of both when assessing the disorder and targeting both in intervention programs. The importance of considering the severity of the symptoms when assessing the disorder would be further discussed below.

Table 2

Autism Spectrum Disorder Severity Specifiers

	Social communication impairment	Patterns of restricted, repetitive behaviours
Level 3: *requiring very substantial support*	Severe deficits in verbal and non-verbal social communication skills cause severe impairments in functioning, very limited initiation of social interactions, and minimal response to social overtures from others. For example, a person with few words of intelligible speech who rarely initiates interaction and, when he or she does, makes unusual approaches to meet needs only and responds to only very direct social approaches.	The inflexibility of behaviour, extreme difficulty coping with change, or other restricted/repetitive behaviours markedly interfere with functioning in all spheres. Great distress/difficulty changing focus or action.
Level 2: *requiring substantial support*	Marked deficits in verbal and non-verbal social communication skills; social impairments apparent even with supports in place; limited initiation of social interactions; and reduced or abnormal responses to social overtures from others. For example, a person who speaks simple sentences, whose interaction is limited to narrow special interests, and who has markedly odd non-verbal communication.	The inflexibility of behaviour, difficulty coping with change or other restricted/repetitive behaviours appear frequently enough to be obvious to the casual observer and interfere with functioning in a variety of contexts. Distress and/or difficulty changing focus or action.
Level 1: *requiring support*	Without supports in place, deficits in social communication cause noticeable impairments. Difficulty initiating social interactions, and clear examples of atypical or unsuccessful response to social overtures from others. May appear to have decreased interest in social interactions. For example, a person who can speak in full sentences and engages in communication but whose to-and-fro conversation with others fails, and whose attempts to make friends are odd and typically unsuccessful.	The inflexibility of behaviour causes significant interference with functioning in one or more contexts. Difficulty switching between activities. Problems of organisation and planning hamper independence.

Note. Adapted from the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-V; pp. 50-51), *American Psychiatric Publishing, a Division of American Psychiatric Association, Arlington, V.S. Copyright O 2013. All right reserved*

1.2. The Hypothesis of the Spectrum of Autism

1.2.1. From a Categorical Towards a Dimensional Conceptualization of the Disorder

Unlike other well-established disorders (e.g., Down Syndrome) the diagnosis of autism has not taken a direct path with several deviations and divergences in its diagnostical history. Yet, if there is something that authors could not agree more is about the high heterogeneity of symptoms presented across individuals diagnosed with ASD. This heterogeneity includes the presence of some ASD core symptoms (Masi et al., 2017), different cognitive profiles (Lombardo et al., 2016), comorbid conditions (Matson & William, 2013), brain atypicality (Floris et al., 2016), and genetics (Betancur, 2011).

Autism was the name granted to this amalgam of symptoms which appeared for the first time in the DSM-I (APA, 1952) as a form of childhood schizophrenia. Thirty years of intensive research drove to the first inclusion of autism as a separate disorder, labelled as infantile autism, in 1980 in the DSM-III. Research continued and, in 1994, *pervasive developmental disorder not otherwise specified* and *Asperger syndrome* was added to the DSM-IV through the expansion of the diagnostic criteria to include subtypes of autism (Baker, 2013).

However, in the last twenty years, autism disorder has walked from a categorical (affected vs. non-affected) to a dimensional approach (behaviours are continuously distributed). To understand this transformation, it is important to draw attention to how the autism disorder was defined in the DSM-IV-TR (APA, 2000). In the previous version of this diagnostic manual, autism took part in the broader category of *pervasive developmental disorders*, where it had a separate label, *autism disorder* which differentiated it from the rest of the disorders included in the nosology and from non-affected populations. Autism disorder could be diagnosed based on three core dysfunctions: qualitative impairment in social interaction, qualitative damage in communication, and restricted, repetitive, and stereotyped patterns of behaviours, interests, and activities. It is worth mentioning that this categorical diagnosing highlights the conjoint presentation of, at least, six of the qualitatively different symptoms included inside each dimension in an individual to provide a formal diagnosis (see Figure 1, Diagram A, p. 44).

Likewise, some authors have suggested that autism is best represented as a discrete entity that differs from normal behaviour in unaffected individuals (Frazier, et al., 2010). Taxometric analyses of some autistic measures, like the Autism Quotient (AQ; Baron-Cohen et al., 2001), supported this categorical structure and observed up to six distinct subtypes of autistic profiles (James et al., 2016). Accordingly, studies have made considerable efforts to define ASD subtypes based either on cognitive profiles (e.g., Georgiades et al., 2013; Marquand et al., 2016) or brain features (Hong et al., 2018). Taxometric approaches could certainly be useful for some fields of study (e.g., to explore different neural subtracts for each autistic subtype), but, in terms of parsimony, considering autism as an aggregation of qualitatively different symptoms could interfere with the composition of the samples of participants as individuals with a diagnosis of ASD (inclusion criteria) would present widely different symptoms and behaviours. Even though samples may be selected with a prior classification of individuals according to their specific symptomatology in practice, it may be difficult to obtain a homogeneous sample and even, if possible, this would cost time and resources.

On the opposite pole, the conceptualization of autism as a continuum would imply that the autistic defining expressions are dimensionally (not categorically) distributed in the general population (Constantino & Charman, 2016) and, so, differences between typicality and autistic behaviours at clinical levels are a matter of degree. The extension of the dimensional approach involves the autism spectrum hypothesis, which proposed a broadening of the spectrum so that, as we have just mentioned, autistic behaviours would be presented not only in clinical populations but also within and throughout the general population (Hoekstra et al., 2008). More important, when expressed as subclinical levels, these behaviours have been reported qualitatively similar to those observed in clinical autism levels (Constantino, 2018). This perspective is supported by multiple studies in which these autistic-like expressions have shown higher intensity among first-degree relatives (including sibling and parents) of diagnosed ASD than that observed in the general population (Constantino et al., 2010; Frazier et al., 2015; Lyall et al., 2014; Page et al., 2016). Furthermore, in higher levels of the clinical impairment, studies at later ages found strong stability of the broad ASD distinction (Stewart et al., 2018), while substantial shifting occurred within specific DSM-IV diagnoses. Similarly, most studies of ASD symptoms and cognitive processes have identified only quantitative but not qualitatively distinctions among DSM-IV disorders.

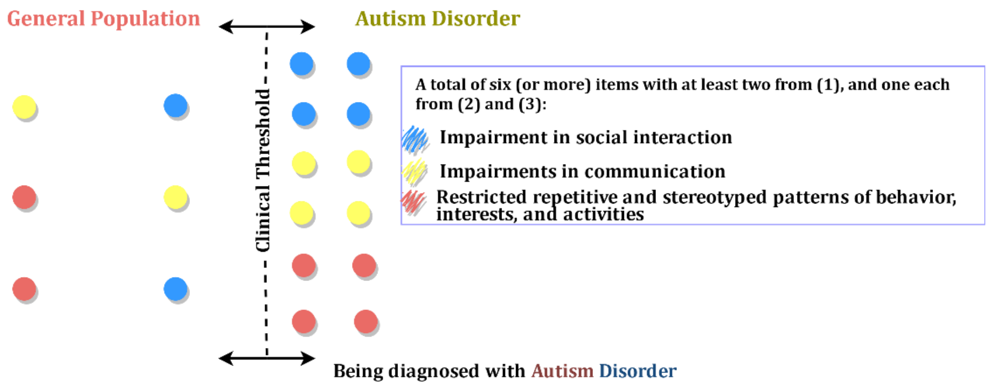
As an intermediate stage in the path towards a dimensional conceptualization, the DSM-V has opted for a hybrid model specifying a categorical distinction of the two dimensions - Social Communication and Interaction (SCI) and Restricted, Repetitive Behaviour (RRB)[3] that have collapsed the three DSM-IV-TR domains together with three levels of severity specifiers for each dimension (dimensional contribution; APA, 2013; see Figure 1, Diagram B). The ASD encompasses the DSM-IV's separate *pervasive developmental disorder* diagnoses: autistic disorder, Asperger's disorder, childhood disintegrative disorder, and pervasive developmental disorder not otherwise specified (APA, 1994) inside the autism spectrum of behaviours. Thus, DSM-V has presented a comprehensive model with a unitary ASD category superimposed on two primary dimensions each scaled in the intensity of their behavioural expressions. The arrival of the ICD-11 will parallel this hybrid diagnosis of ASD as it has been advanced (WHO, 2019).

[3] ASD is diagnosed upon the observations of all the behaviours in the domain of persistent deficits in social communication and social interaction and at least two from the restricted, repetitive patterns of behaviour, interests, or activities one.

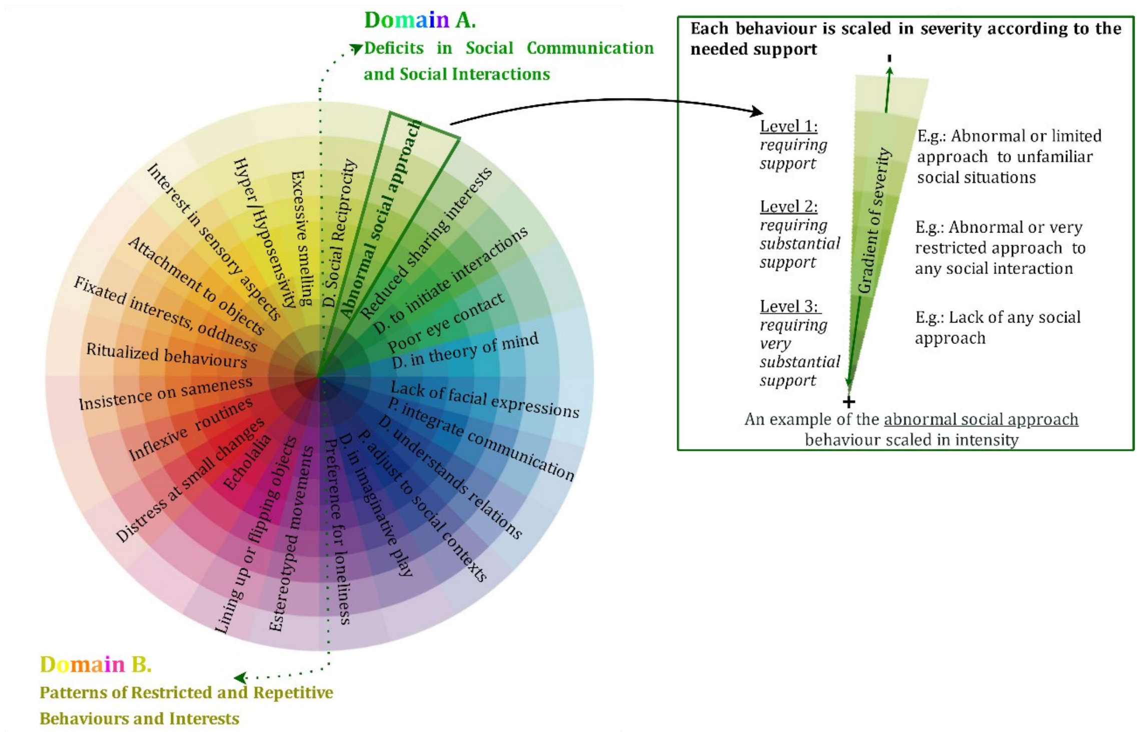
Figure 1

Categorical and Hybrid Theories of Autism in the DSM-IV-RT and DSM-V

A. The Categorical Conceptualization of the Autism Disorder (DSM-IV-RT): An aggregation of symptoms until meeting the clinical threshold



B. The Hybrid Conceptualization of the Autism Spectrum Disorder (DSM-V): A spectrum of varied behaviours organized around two core domains



Note. Diagram A shows the conceptualization of the autism disorder as a grouping of six or more items, the maximum number of behaviours or the severity of them is not specified during diagnosis. When a person presented less than six indicators, they did not reach the clinical threshold but autism was suspected, a pervasive developmental disorder not otherwise specified was diagnosed. Diagram B shows a hybrid model where all the socially impaired behaviours must be presented together with at least two of the restricted and repetitive ones for considering the diagnosis. The differences in the patterns of restricted and repetitive behaviours and other associated characteristics (e.g., intellectual disability) have enabled the inclusion of all the broader categories of *pervasive developmental disorders* inside the autism spectrum. Both dimensions of ASD have been scaled in severity according to the level of functionality of the person. An example of a behaviour graded in severity (abnormal social approach) is provided. D. means deficits; P. means problems.

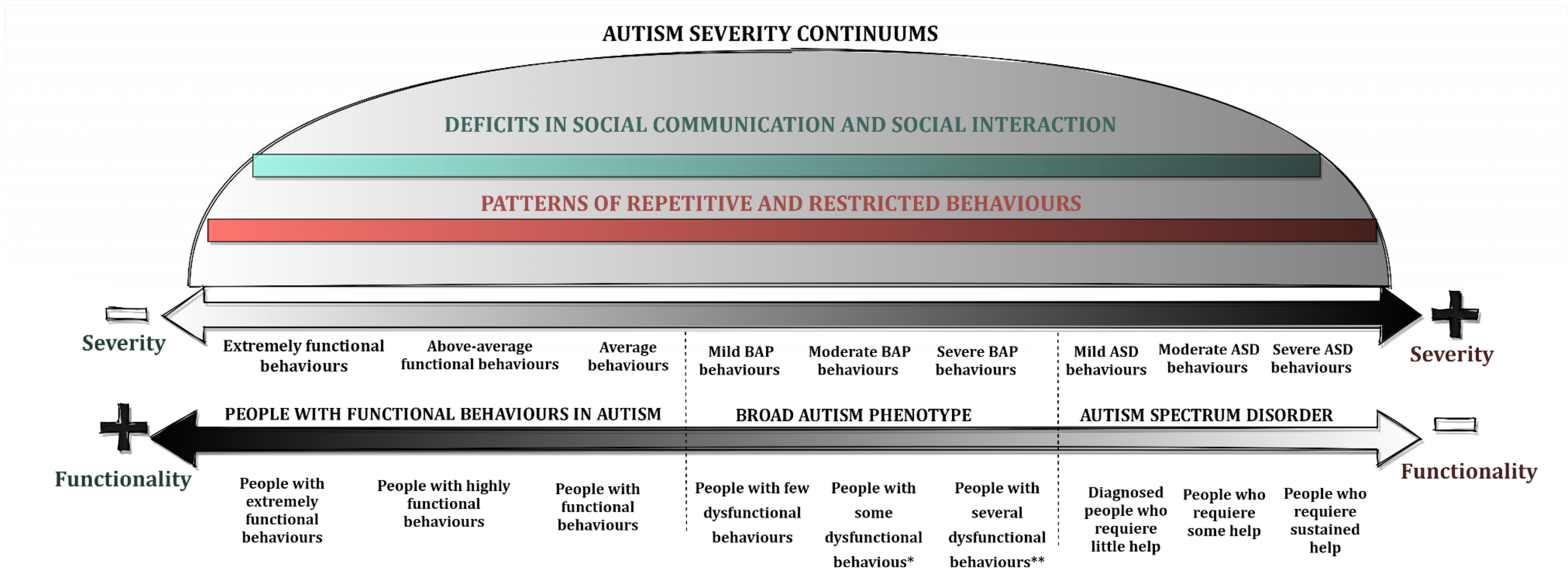
1.2.2. Implications of Spreading the Autism-related Behaviours Beyond the Clinical Relevance, Towards and Across the Vast General Population

Throughout these pages, I have taken a step forward and I have considered a fully dimensional approach to autism. As it is shown in Figure 2, this dimensional perception implies that autistic associated behaviours are disposed throughout one of two continuums of gradual magnitude instead of representing an amalgam of behaviours. Individuals could behave in varying degrees of severity inside both main dimensions: difficulties in social communication and social interaction (first dimension), and restricted and repetitive patterns of behaviours, preference for sameness and routines, and sensory abnormalities (second dimension), according to their level of functional independence (De Groot & Van Strien, 2017). Importantly, these dimensions correspond to both domains proposed in the DSM-V and only those behaviours that are capable of being scaled inside each of both dimensions and that have been included as prototypical autistic expressions in the DSM-V are considered relevant for autism. This simplifies the autism conceptualization as I present what I have called the “spinal columns” of autism.

According to the dimensional hypothesis, differences in expressions along the general population are presumably due to divergences in how the same behaviour is exhibited at different levels of severity. Of great relevance for the present dissertation, this extension of the autistic into the general population comprehends the broad autism phenotype (BAP). The BAP, a term that will be deeply explained later, is defined as the milder expressions of autistic characteristics (Hurley et al., 2007). That is, ASD expressions at community levels (De Groot & Van Strien, 2017). This idea links both the BAP and the ASD implying that every single individual (and behaviour) could be inserted somewhere on the autistic spectrum and the amount of autism one exhibits is hypothesised to be correlated with several domains of functioning, such as cognitive, motor, and social one (De Groot & Van Strien, 2017). Finally, as ASD could lead to thinking exclusively in severe and dysfunctional expressions at clinical levels and considering that this dissertation consists of autistic behaviours at BAP levels (milder, less severe levels of expressions), the concept of autism is going to be used through these pages to express the link between these behaviours that could be scaled throughout the two continuous previously mentioned. Equally, it also enables the consideration of both ASD (clinically relevant levels of expression) and BAP inside the same continuum.

Figure 2

An Expansion of the ASD Model (APA, 2013) Towards Lower Unclinical Levels of Expression



Note. An expansion of the autism continuum towards lower unclinical severity levels has been proposed from a dimensional perspective. The diagram represents the two severity continuums graded from less to more intensity (brighter to darker pink or green). Both continuums are included inside the *autism sphere* as they may concur in autism. The intensity of the expressions is related to the level of functionality of the person. Less severe behaviours correspond to a more functional individual. The more intense the expression of behaviour, the less functional the person is. When a person behaves intensively compromising him/her functional independence in daily life, this person is diagnosed with ASD. The people above in the continuums may show similar behaviours without compromising their functionality. * This group includes those with distant relatives diagnosed with ASD. ** This group includes those with close relatives diagnosed with ASD.

1.3. Applications of Characterising Autism as Continuums of Gradual Severity

Myriads are the benefits of exploring autism from a dimensional perspective. As such, studies have highlighted its usefulness for the traceability of the aetiology of the disorder (Constantino and Todd, 2003; Le Couteur et al., 1996; Losh et al., 2009; Sung et al., 2005). Regarding this issue, it must be noted that an endophenotype comprehends several genetic markers associated with a specific factor that can be expressed by the carrier (Gottesman & Gould, 2003). The endophenotype eases the study of the genetic components of the disorder and its relationship with the intensity to which its expressions are disposed in the general population. As an example, large genetic-epidemiologic studies have found that heritable susceptibilities to some subclinical patterns observed in the general population are near-completely overlapped with the genetic underpinnings of clinically relevant autistic levels (Robinson et al., 2012). Among these subclinical patterns, the BAP corresponds to the autistic endophenotype, as its name indicates. Originally, the discovery of the BAP was used to study the liability of the ASD since relatives of people diagnosed with ASD (especially first-degree relatives such as twins, siblings, and parents) presumably had a higher chance of presenting some autistic-like expressions and even being clinically impaired (for a review, see Bailey et al., 1998).

On the other hand, the study of the BAP has also been applied to the research of personality, behavioural, and neuropsychological profiles in parents and siblings of people diagnosed with ASD (Klusek et al., 2014; Losh et al., 2009; Ruzich et al., 2015) and, later on, in the general population. The utility of this kind of research consists in the study of several psychological processes that have been observed to be altered in individuals diagnosed with ASD (e.g., social cognition) through BAP samples, as it has been found that they correlated with certain BAP behavioural components (D'Cruz et al., 2013; Landry & Chouinard, 2016).

In addition, the dimensional perspective implies that typically developed individuals may express autistic behaviours less intensely, providing the opportunity to explore how each specific behaviour is related to ASD and their expressions at different intensities. This is highly relevant for the discussion of the intrinsic link of each dimension as well as its contribution to the whole disorder. Another advantage of expanding the continuums lies in conducting studies in broader and more heterogeneous samples (Landry & Chouinard, 2016). This means a better representation of the population and a guarantee of higher statistical power (Mitchell & Jolley, 2013).

Methodologically speaking, these studies broaden the possibilities in terms of the methods applied, tasks used, experimental environments, and so on (i.e., it could be easier to conduct electrophysiology studies in average functional people who present moderate to high BAP-levels than in people diagnosed with ASD). It also increases control guarantees since we eliminate the presence of common comorbid disorders in clinically-relevant levels of autism (ASD; Simonoff et al., 2008, such as Intellectual Disability) and we can pair participants according to their chronological and mental age and cognitive development since it probably corresponds to the average one (Elsabbagh et al., 2012).

Finally, as it has been mentioned above, we cannot disregard the advantages of improving the knowledge about how psychological processes are related to autism through average functional samples and how this could serve to polish intervention programs and, ultimately, improve the quality of life of individuals located in higher and clinical levels of the continuums. For example, Ingersoll and Hambrick (2011) found that anxiety and depression feelings, as well as the perception of lack of social support, would be mediated by the severity of the autistic expressions in individuals who had shown BAP levels. As in people diagnosed with ASD, it has been reported that people at BAP levels of severity feel lonelier and have significantly fewer friends or they are more superficial (Jamil et al., 2017). Those studies support the idea that a better comprehension of the lower levels of autism would help us design better intervention programs that would have a higher impact in their quality of live. Likewise, itself would also improve the quality of life of people who exhibit behaviours at BAP levels and still do not perfectly understand the way they interact with others, their likes, and/or preferences. In the same way, working, for instance, on anxiety and depression feelings on parents of children diagnosed with ASD will mediate parental care and eventually will impact the quality of life of their children.

1.4. Contributions of Both Genetic and Idiosyncratic Development in the Understanding of How Autism Expresses Differently across Countries

Despite being a very prolific field of research, the background of autism is only partially understood. Authors have suggested that it has multiple aetiological factors, both genetic and environmental and they have made considerable efforts in trying to find the genetic traceability for autism. The study of its phenotype, the BAP, gains meaning in this sense. Additionally, the gene-environment interaction should also be studied as it can lead to epigenetic abnormalities and cause alterations in the brain anatomy and connectivity characteristic for autism (Schaevitz & Berger-Sweeney, 2012). All the findings provided from each branch of research have driven to create a heterogeneous disorder both in the behaviours that individuals may express and in their aetiology.

From the branch of the genetic liability of autism, it has been suggested that (i) there is little variation in the expressions of autism and (ii) core behaviours are consistent across cultures, ethnicity, and social class (Cuccaro et al., 1996). In this regard, while most of the studies have been and are still being conducted in Western countries (Abubakar et al., 2016; de Leeuw et al., 2020; de Vries, 2016; Durkin et al., 2015; Elsabbagh et al., 2012; Hahler & Elsabbagh, 2014), interest in genetic studies of autism have emerged and increased in different countries, as shown by recent studies from Hong Kong, Taiwan, and China (Chien et al., 2011; Siu et al., 2016; Wang et al., 2016). Interestingly, findings have pointed in the same direction and have suggested that candidate risk genes to autism identified in European cohorts could be the same as those documented in other cohorts like Chinese ones (Warrier et al., 2015).

In the same way, despite having a complex multifactorial aetiology, twin studies proved a strong genetic contribution to autism. As such, the rate of higher levels of autism (reaching the diagnosing of ASD) in monozygotic twins is 70–90%, decreasing around 30% in dizygotic twins (Rosenberg et al., 2009; Hallmayer et al., 2011; Ronald & Hoekstra, 2014) and to 3–19% in siblings in general (Ozonoff et al., 2011; Constantino et al., 2013). The recurrence risk for presenting clinically impaired levels of autism in the general population is estimated at 1%. This recurrence rate increases around 10-15% with one previously affected sibling and around 50% with two affected siblings in the family (Ozonoff et al., 2011).

Nowadays, the genetic aetiology of the disorder is recognized in ~ 25–35% (Wiśniowiecka-Kowalnik & Nowakowska, 2019). By now, more than 800 genes have been included in the AutDB (Human Gene Module - SFARI Gene), a database of genes implicated in autism (Basu et al., 2009), but the strength of the evidence supporting each of these observations varies greatly and for many of these studies the effect sizes are quite small.

If we could draw a conclusion after exposing the prolific research in genetics, it would be that autism is gaining interest around the world. Epidemiological findings have been reported from Europe and South America (Montiel-Nava & Peña, 2008; Oliveira et al., 2007). Likewise, diverse countries such as Taiwan (Lung et al., 2010), Malaysia (Kasmini & Zasmani, 1995), Zimbabwe (Khan & Hombarume, 1996), United Arab Emirates (Sartawi, 1999), Portugal (Oliveira et al., 2007), Brazil (Bandim et al., 2002) and some others have explored autistic-like behaviours. Diagnostic criteria have been widely accepted in many countries (Amaral et al., 2011; Cohen & Volkmar, 1997). However, when parents are reported that their child has been diagnosed with ASD, they could form different ideas and thoughts depending on the country they belong to.

Despite being a universal disorder with strong biological underpinnings that occurs with similar core behaviours, autistic expressions appear to be susceptible to cultural influences (Matson et al., 2017). Cultural idiosyncrasies could play an important role not only in social acceptance, understanding, and support for different populations but also it could be possible that certain autistic behaviours would be interpreted differently across cultures. As such, lack of eye contact is viewed as unusual in Western countries while it could be considered polite and respectful in China and South Korea (Volkmar et al., 2014). However, other behaviours, such as deficits in communication are considered problematic in all cultures (Volkmar et al., 2014). The differential understanding and acceptance of certain characteristics could increase or decrease their expression inside a specific culture and not in others. Thus, it could be possible that culture modulates the autistic behaviours that individuals are presenting.

Even though the role of culture seems clear, its influence in the developmental trajectories of autism has been largely ignored (Daley, 2002) until recently where this issue is gaining more and more attention (see Elsabbagh et al., 2012, for a review). This lack of interest has been feeding by misheld beliefs that autism is a "common Western cultures disorder" summed to its consideration as a neurobiological disorder with a stable set of expressions all over the world. However, understanding cross-cultural expressions of autism is especially of relevance since cultural judgments about correct and/or typical behaviours impact directly the way parents, caregivers, and the population, in general, conceive autism (Daley, 2002; Mandell & Novak, 2005). According to Rogler's framework (1993), cultural norms mediate the endorsement of behaviours, the severity of their expression, and their rating in measurement tools.

Also, the differential cross-cultural conception of autism could compromise the perception of some or other behaviours, their consideration as relevant expressions, and how severe they are thought, that is, the way autism is assessed by professionals. In this regard, Rogler (1993) proposed a general hierarchical three-level framework for understanding how culture[4] could influence psychological assessment. This hierarchy could be applied to the assessment and diagnosis of ASD. Specifically, the first level entails how culture impacts the assessment of specific expressions and their severity. The second level corresponds to how culture hinders the configuration of behaviours into disorders, which is particularly relevant to the use of culturally sensitive diagnostic and screening measures (e.g., self or informant-report questionnaires). Finally, at the third level, both individuals involved in the diagnostic process (the therapist and the patient) would bring their personalities, cultural influences, and idiosyncrasies to the diagnostic situation. Therefore, these cultural factors would be presented in every diagnostic situation modulating the relationship between the person and the psychologist (or the person in charge of the diagnostic procedure).

Equally of importance, available assessment tools for children are mostly parent or teacher reports. Thus, how parents and teachers describe their child's behaving and its severity may be influenced by the aspects that are considered more worrying within the culture. This has been documented in some studies. For example, research carried out in the United States have suggested that American

[4] Culture can be better operationalised as "*a set of behavioural norms, meanings, and values or reference points utilised by members of a particular society to construct their unique view of the world, and ascertain their identity*" (Alarcón, 2009).

parents tend to be more concerned about language delays (Coonrod & Stone, 2004) while Indian parents would tend to have early concerns about social difficulties (Daley, 2004), and Latina mothers may be worried about their child temperament (Ratto et al., 2015).

On the other hand, some parents could camouflage, cover, or even deny socially undesirable behaviours of their children that could lead to societal stigma. This could also be different across cultures (Chung et al., 2012) suggesting that ethnically-based cultural norms would modulate the perception of socially undesirable mental characteristics, and, ultimately, the children endorsement of those symptoms (Matson et al., 2017). In this regard, researchers have found some international differences in the endorsement of deficits related to ASD, including core behaviours (Freeth et al., 2013; Matson et al., 2011), challenging ones (Chung et al., 2012), social skills (Sipes et al., 2012), and sensory issues (Caron et al., 2012).

Even though authors have documented plenty of those cultural influences during the diagnostic procedure, they did not reach an agreement on their magnitude and, also, there does not exist any predictive model. Culture represents a complex and pervasive construct and, as such, it could be really difficult to determine the magnitude of the cultural influences in a disorder, its development, assessment, diagnosis, and intervention. Nevertheless, this complexity does not imply disregard the role of the culture because this would inevitably guide to simplistic models and theories of the psychopathological disorders. Instead of trying to control cultural influences, they should be recognised as essential aspects of all phenomenological experience and, thus, factors that should be integrated with every psychological disorder including ASD.

Second Part. The Broad Autism Phenotype: Filling the Gaps in the Autism Spectrum and Enhancing the Understanding of the Distribution of Autism across a Specific Population

The extension of the autism severity continuums towards the general population includes the BAP. This phenotype has been described as a set of milder autistic expressions qualitatively similar to those observed at higher clinical levels (Hurley et al., 2007) which are continuously spread in the general population (Constantino & Todd, 2003; Hoekstra et al., 2007; Hurst et al., 2007; Stewart & Austin, 2009). This dimensional approach is supported by numerous authors (e.g., Morrison et al., 2018; Nayar et al., 2021; Pruit et al., 2018; Wainer et al., 2011) who have reported that behaviours at BAP levels are the same as the primary characteristics of ASD (grouped into two dimensions, A. Social Communication and Social Relation Impairment and B. Patterns of Restricted Repetitive Behaviours and Interests; APA, 2013). The second part of this chapter and successive ones will be focused on less severe autistic expressions like those observed at BAP levels of the continuums.

2.1 The presentation of the BAP: From its First Discovery to its Formal Definition Based on Empirical Evidence

The broad autism phenotype (BAP) was first reported by Leo Kanner (1943), who observed that the parents of children diagnosed with autism presented subtle expressions of autistic-like traits such as an obsession with details, social awkwardness, and rigid behaviours. Further studies agreed to indicate that relatives (mainly twins, parents, and siblings) of those individuals are more likely to express mild autistic behaviours, regardless of diagnosis (for a review, see Bailey et al., 1998). Finally, gene studies helped to shed light on this line by suggesting that ASD is a heritable disorder and those generic variants for autism when expressed as milder but qualitatively similar forms, constitute the BAP (Piven, 2001; Szatmari et al., 2007).

In Piven et al. (1997a, b) studies, authors conducted more formal investigations in parents of children diagnosed with ASD and identified three core deficits that established the foundations for the upcoming operational definition of the BAP. Taking these findings, Hurley et al. (2007) finally defined the BAP as a set of milder personality characteristics and language deficits clustered around three main domains, namely, aloof personality, rigid personality, and pragmatic language problems.

According to Hurley et al. (2007, p. 1681), an aloof personality referred to limited interest in or enjoyment of social interaction; a rigid personality was related to resistance, and/or difficulty adapting to change; and, finally, pragmatic language problems comprised those deficits in the social use of language that led to difficulties with effective communication and/or conversational reciprocity. Further studies also reported that the aloof personality involved a lack of social responsiveness, reduced social abilities, and scarce social engagement (Constantino et al., 2006; Wheelwright et al., 2010; Whitehouse et al., 2010); the rigid personality was manifested in behavioural rigidity, a tendency toward perfectionism, stubbornness, and stereotyped behaviours (Losh et al., 2008; Murphy et al., 2000; Narayan et al., 1990); and pragmatic language impairment was related to supra-linguistic aspects such as problems in respecting turn-taking in speaking, becoming side-tracked in conversations, and difficulties in maintaining the topic of conversations (Seidman et al., 2012).

It is worth noting that, with this definition, the BAP conceptualization was paralleled to that proposed at that moment for the autism disorder (AS) by the DSM–IV-TR (APA, 2000). Yet, the BAP operationalization should have been revised in 2013 with the arrival of the DSM-V and the conceptual change of autism disorder to ASD (see first part of the present dissertation). With this revision, the BAP should have been changed to comprehend only two dimensions as the ASD now does. However, eight years later, this is still a to-do task.

2.1.1 Evidence Congruent with the Equality of BAP and ASD Expressions

Agreeing with the idea of a continuum of gradual severity, expressions of the BAP have also been observed throughout the general population (Constantino & Todd, 2003; Frazier et al., 2015; Happé et al., 2006; Jamil et al., 2017; Sasson et al., 2013c), aggregating its incidence in first-degree relatives of individuals diagnosed with ASD (Bailey et al., 1995; Gerdts, 2013; Losh et al., 2010; Piven et al., 1997a, b; Sung et al., 2005). This prevalence is even higher in multiple incidence case families (Bernier et al., 2012; Losh et al., 2008). In this regard, while authors have stated the prevalence of BAP severity levels among families where relatives exhibited autism behaviours ranging from 12 to 30% depending on the incidence rate in each family (Bishop et al., 2004; Hurley et al., 2007; Sasson et al., 2013a), Sasson et al. (2013a) study has documented that around the 8% of the participants without a history of relatives diagnosed with ASD overtook the Broad Autism Phenotype Questionnaire's (BAPQ, Hurley, et al., 2007) cutoffs. These findings could be also summed as support for the genetic (Devlin & Scherer, 2012) and heritable (Colvert et al., 2015) nature of the ASD since biological relatives of people diagnosed with ASD have shown higher severity of BAP behaviours. Finally, those BAP expressions seem to persist into adulthood representing stable preferences (Howlin et al., 2015) and explaining how autistic genes could be transmitted from parents to newborns.

Regarding specific manifestations, many studies have shown similar dysfunctional behaviours, but milder, than those related to ASD in typically developing individuals (De Groot & Van Strien, 2017). Inside the dimension of social communication and interaction (Figure 2), authors have found a reduced quantity and quality of friendships (Jamil et al., 2017; Piven et al., 1997b) with lower numbers and/or length of friendships, lower relationship satisfaction, and poorer interpersonal interactions (Faso et al., 2016; Jobe and White, 2007; Wainer et al., 2013); an aloof personality (Austin, 2005; Murphy et al., 2000; Piven et al., 1997b); deficits in the social use of the language (Landa et al., 1992; Piven et al., 1997b); weak central coherence (Happé et al., 2001), reduced emotional understanding (Szatmari et al., 2008); abnormal response to gaze (Adolphs et al., 2008; Scheeren & Stauder, 2008); impairment in executive function (Hughes et al., 1997; Ozonoff et al., 1993) and social cognition (Baron-Cohen & Hammer, 1997; Losh & Piven, 2007; Sasson et al., 2013c; Tsang et al., 2016). Concerning the restricted and repetitive patterns of behaviours dimension, a smaller number of studies have been performed. Authors have confirmed that BAP is related to little interest in novelty, resistance to change, and strong attachment to routines (Bernier et al., 2012; Bishop et al., 2004; Hurley et al., 2007; Losh et al., 2008). As far as we know, no study has included stereotyped and repetitive motor mannerism when studying the BAP (see, Cruz et al., 2013, for a review). In my view, all the abovementioned studies have supported and claimed for the updating of the BAP operationalization following the ASD one in the DSM-V (APA, 2013). Even though those studies have assumed that

BAP and ASD behaviours are the same, in essence, with different severity, consensus needs to be reached in the conceptualization of the BAP to continue conducting research. This consensus would come to reinsert the BAP in both dimensions of the autism severity continuums and to define how autistic behaviours express themselves at BAP levels. Besides the highlights of prototypical diagnostic deficits and personality preferences, research has also drawn attention to the BAP neuropsychological profile. For example, equally to diagnosed individuals, typically developed ones scoring high in the Autism Quotient (AQ; Baron-Cohen et al., 2001) were faster and more accurate on tasks measuring visuospatial ability (Cribb et al., 2016; DiCriscio & Troiani, 2018; Grinter et al., 2009; Stewart et al., 2009). As another example, both people diagnosed with ASD and their parents with high BAP severity performed significantly worse in measures of social cognition such as inferring psychological states from viewing slices depicting only the eye region of faces; identifying emotions from gradual facial expressions; stating trustworthiness judgments based on emotional expressions; and so on (Losh et al., 2009). Finally, a study conducted in elder adults (aged over 60 years) who have overtaken the BAPQ cut-offs demonstrated poorer executive functioning and episodic memory (Stewart et al., 2018).

Equally, comorbidity of BAP has documented similar mental states to those found in ASD such as higher depressive feelings and greater negative affectivity in mothers of children diagnosed with ASD (Barker et al., 2011; Ingersoll et al., 2011; Ingersoll & Hambrick, 2011); lower level of reported social support (Ingersoll & Hambrick, 2011); anxiety, depression, and obsessive-compulsive disorder; and other cognitive deficits like weak central coherence and neurological processing (Gerdts & Bernier 2011; Sucksmith et al. 2011).

However, not all findings are supportive of ASD and BAP equality. Especially, studies focusing on the link between executive functioning and autistic-like expressions in typically developed individuals are inconsistent (Kunihira et al., 2006; Maes et al., 2013; Ridley et al., 2011). In the same way, even though pragmatic language impairment may be seen as a reliable characteristic of the BAP, other language deficits, such as delays and difficulties in reading and writing acquisition and difficulties in structural language use, although present in some of the parents of children diagnosed with ASD, have not been confirmed in other studies (Lindgren et al., 2009; Whitehouse et al., 2007, 2010). Nevertheless, a recent review has called attention to the problem with these studies; they examined very specific behaviours and from a categorical view (upper and lower extremes of the continuum; De Groot & Van Strien, 2017). By taking a dimensional approach that includes a wider range of expressions in social functioning and including the natural range of autistic scores it could be possible to amend these discordances. According to this, the findings of De Groot & Van Strien (2017) supported the dimensional conceptualization of the BAP by showing a continuous measure of autistic manifestations related to a continuous measure of social functioning.

2.2 The State of Art: What More Recent Literature Says about BAP and ASD Equality of Definitions

To make accurate scientific observations, researchers need to determine what we want to observe and how to make this observation. First, we need to conceptualise the target of interest. A conceptualization of the BAP as the subclinical expressions of the ASD has been already proposed and accepted by the broad scientific field (e.g., Hurley et al., 2007; Rubenstein and Chawla, 2018; Sasson et al., 2013c). The second step should be to operationalise the BAP in a specific and identifiable set of behaviours that would enable its subsequent measurement. Lastly, the operationalisation of the construct is depicted in specific indicators represented in the items of a measurement instrument.

Concerning the second step, the BAP has been extensively documented and studied since it was firstly reported. Although this is now a more mature field, there is still no formal conceptualization of the BAP except that the one proposed by Hurley et al. (2007). Most of the authors have accepted and cited Hurley's definition of the BAP when they have defined this phenotype in their studies. For example, in a recent cross-sectional study, Hirokawa et al. (2020) defined it as a group of key components such as aloofness, rigidity, and pragmatic language difficulties referring to them as Hurley et al. (2007) originally did. This implies that the BAP has barely changed since it was first formulated. However, as above stated, several important changes have occurred in the ASD operationalisation with the arrival of the DSM-V (APA, 2013), later supported by the forthcoming ICD-11 (WHO, 2019). In both manuals, the ASD is now considered a dimensional disorder comprising only two core domains and which varies widely in the intensity of its expressions, the intellectual level, and the functional disability of each specific person (Geshwind, 2009) remarking its extension beyond the clinical threshold (Landy & Chouinard, 2016) and implying, therefore, an alignment with the BAP definition as the ASD phenotype. Nevertheless, the updating of the ASD definition has not even been accompanied by a revision of the BAP. This has generated a discrepancy between both conceptualizations that needs to be addressed because studies are comparing both populations referring to two uneven conceptualizations detached in the time. This could, in turn, lead to inaccurate conclusions about their results.

Regarding this issue, an emerging number of authors have passed to directly associate the BAP with the non-clinical expressions of the ASD and listing some examples of typical behaviours. In a recent systematic review, Rubenstein and Chawla (2018) defined the BAP as "a collection of sub-diagnostic autistic traits" which include, among others, pragmatic deficits, communication difficulties, poor social abilities, rigidity, and stereotyped behaviours. This practical approach followed by some authors has let us explore BAP-ASD deficits from a quantitative perspective and has contributed to the knowledge of the specific autistic behaviours and their expression and distribution in clinical populations, families, and even the general population without neglecting the fact that both constructs have different conceptualizations.

Other authors have tried to revise this issue by comprising the three BAPQ components proposed by Hurley et al. (2007) into two, conceptualized as both social and non-social expressions of the BAP, where the social component has included both social impairment in social relations and social communication (i.e., aloof personality, pragmatic language impairment), and the non-social one has been constituted by a rigid personality (Morrison et al., 2018; Sasson et al., 2013b; Taylor et al., 2014). Nonetheless, this way of collapsing the BAP structure without taking into consideration the ASD definition could also mislead research conclusions. As such, ASD now contemplates some stereotyped language behaviours like rituals, greetings and goodbyes, topic control, and maintenance (Russell & Grizzle, 2008), together with echolalia and idiosyncratic phrases as representative behaviours of the restricted and repetitive behaviours and interests' domain (APA, 2013). For this reason, it may be possible that the Pragmatic Language component from the BAPQ

(Hurley et al., 2007) could contemplate not only prototypical social deficits but also restricted and repetitive behaviours. Therefore, the studies that collapse aloof and pragmatic language components from the BAPQ in what it is called “social BAP” (Sasson et al., 2013c) should be interpreted with caution.

Considering all the abovementioned, it seems clear that the BAP operationalisation should be reviewed and aligned with the current ASD one. In the view of the present dissertation, the field of research in autism urgently needs this revision in the BAP construct as its exploration continues to be very prolific nowadays.

2.3 Justifying the Need: Advantages of Linking the BAP and the ASD

Using the BAP as a landmark to explore autism may bring several potential benefits. From a dimensional perspective, research conducted in populations showing BAP severity levels is important as it can provide a further understanding of the continuum of autism itself and, also, increases the information regarding how differential severity of core ASD domains is related to various abilities and disabilities (e.g., interpersonal functioning) at BAP levels in particular (Jamil et al., 2017; Pruitt et al., 2018). Here, I will focus on the improvements that the study of the BAP could provide in the assessment, diagnosis, and intervention of ASD. Secondly, the theoretical implications of studying the BAP for increasing the knowledge about the aetiology of autism will be also indicated. Finally, the empirical advantages of studying autistic-like behaviours in these samples instead of at clinical levels must be necessarily mentioned as they are going to represent the foundation stone of this dissertation.

2.3.1 Practical Implications: Assessment and Intervention in Clinical Populations

From a practical view, the exploration of autistic-like behaviours in broader and more heterogeneous samples could contribute to gain a better understanding of the disorder, its domains, subdomains, and prototypical behaviours. I consider that this may have a direct impact in the assessment process and the subsequent intervention practices. Furthermore, it would also be beneficial in the development of new tools for evaluating and diagnosing ASD. Professionals could equally try new techniques and interventions at non-clinical levels before applying them to the clinical ones.

Regarding the assessment of the autistic-like behaviours, it must be mentioned that some tools applied to populations presenting BAP severity levels were originally designed to characterise autism in adults with standard intelligence. Researchers using these tools (the AQ or the Social Responsiveness Scale -SRS-II; Constantino & Gruber, 2012) have reported that adults with a diagnosis of ASD tend to score higher than those with BAP levels. In a large meta-analysis, Ruzich et al. (2015) indicated that individuals with a diagnosis of ASD obtained average scores of 35 on the AQ, which represents more than twice the general population's mean score (17). Similarly, individuals diagnosed with ASD got mean raw scores greater than 80 in the SRS, while the mean raw scores for participants at typical levels of severity were generally located below 40 (Constantino & Gruber, 2012). These results have supported the idea of diverse degree of severity in the autism expression. By contrary, it might be noted that BAP measures should not be used for establishing an ASD diagnosis, but they could be used for exploring both the similarities and the divergences in individuals at BAP and ASD levels with standard intelligence in their adulthood with and without a diagnosis of ASD.

Finally, I would like to provide an example of how studies conducted in people at BAP levels of severity may complete studies addressing ASD levels by throwing light on controversial results or which cognitive processes would be interesting to study in relation to autism. As an example, there are a growing number of studies that have their interest in the cognitive state of elder people diagnosed with ASD. However, studies have varied in the characteristics of their samples (e.g., young vs. elder adults; adults with ASD severity levels and normal ranged IQ vs. low IQ; severely affected vs. less affected individuals) and the aspects of cognition they examined, so that the pattern of results has resulted inconsistent across them (Davids et al., 2016; Lever et al., 2015; Powell et al., 2017). Interestingly, it has been observed impaired executive functions (e.g., planning and flexibility) in parents and middle-aged siblings of children diagnosed with ASD and in older populations at BAP severity levels (Stewart et al., 2018; Wallace et al., 2016) finding, also, that the latter have reported higher levels of depression and anxiety (Stewart et al., 2018). Even though the number of studies is still scarce, these results suggest that BAP traits could be associated with the aggravation of normative age-related cognitive declines (Stewart et al., 2018). If so, studying the detriment of cognitive processes in later life stages at BAP severity levels could help to explore how clinical ones are expressed in such stages and/or whether they imply an additional risk factor for normal cognitive declines. In extension, these kinds of studies may also lay the foundations for developing intervention programs for adults diagnosed with ASD or those who do not reach the levels for receiving a formal diagnosis but still present high autistic severity which is hindering their lives and impoverishing the key aspects of them.

2.3.2 Implications for Research: Studying the Liability of Autism from a Dimensional Perspective

What can studies conforming the BAP and ASD to one only construct add to the field of research? Firstly, they could contribute to clarify the outcomes of ASD. Let's consider a skill pattern that is uniquely observed at clinical levels, such as low intellectual coefficient (IQ; Szatmari et al., 1995). IQ is a comorbid component that, when presented, aggravates the severity of the autistic behaviours and highly reduces the functionality of the person. This skill pattern that could mark a difference between BAP and ASD levels may represent a milestone in the understanding of ASD and its severity. Now imagine siblings who have obtained an elevated score on a BAP measure, but only one of them has received a diagnosis of ASD. Exploring both qualitative and quantitative differences between these individuals may help us to understand what contributes uniquely to ASD and how these components may be deteriorating, even more, individual independence.

In addition to this, exploring the presence of autistic behaviours in parents of children diagnosed with ASD is essential since genetics, as I have previously mentioned, can be playing a significant role in ASD aetiology (Gaugler et al., 2014). Observing autistic behaviours in a parent of a child diagnosed with ASD suggests a *genetic mechanism* (Robinson et al., 2011). Therefore, clustering the samples according to their familiar predisposition to ASD in research studies could allow statistical control of genotypic influences and possibly phenotypic heterogeneity (Gerdts & Bernier 2011; Sucksmith et al., 2011).

Finally, taking the BAP as a landmark brings the opportunity to examine the developmental paths and correlates of skills involved in ASD in greater isolation. In contrast to ASD (clinical levels of autistic behaviours), which implies multiple deficits from both domains, community people can show elevated scores in BAP measures in only one of the domains as a consequence of the natural distribution of behaviours in the general population—for example, having a rigid temperament does not necessarily imply deficits in social relations. By tracking how autistic-related behaviours are distributed throughout the population, researchers have the opportunity to examine what the delimitation of each autistic manifestation at BAP levels might talk about ASD. This acquires huge importance in the discussion of the intrinsic associations between both ASD dimensions (APA, 2013). As such, twin studies have demonstrated convergent evidence pointing towards the heritability of autistic behaviours inside each dimension (i.e., A. Social Communication Impairment and B. Patterns of restricted, repetitive behaviours), but not across dimensions, together with patterns of correlated and/or uncorrelated impaired skill groups (Brunsdon et al., 2015; Brunsdon & Happé, 2014). These findings suggest that there is not a unique cause that explains autistic deficits in both dimensions, but rather several and multicausal components. Consequently, within the dimensional model of autism, dimensions might be ideally tested in isolation by examining patterns of high and low autistic behaviours independently.

2.3.3 Empirical Implications: Variability of Expressions, Accessibility to the Sample, and Statistical Power

From an empirical perspective, conducting studies with samples at BAP levels enables the investigators to have access to many more participants than in most studies focused on individuals at high ASD levels. Let's think that for each child diagnosed with ASD, there are at least two adults, their parents, who could express BAP severity levels. In many cases, those children will also count on siblings who will amplify the possibilities of studying the divergences in autism expressions along the continuum in more homogeneous samples as they would be about at the same age range.

Similarly, we can think that as BAP severity levels have been documented in twins, siblings, parents, or relatives in general (for a review, see Bailey et al., 1998), families without a history of ASD severity ones, undergraduates, and the general population (e.g., Wainer et al., 2011), studying this phenotype could offer a great range of variability and could provide an enormous richness in the characterisation of the continuum which is, ulteriorly, the purpose of the hypothesis of the dimension of autism. On the contrary, dichotomization[5] can lead to substantive and statistical problems (MacCallum et al., 2002). On one hand, some selection procedures such as dichotomizing autism which is assumed to be continuously distributed in the population according to extreme test scores may bias and reduce the information only to extreme behaviours (e.g., extremely functional vs. extremely dysfunctional people). Instead of only comparing two groups with the most extreme scores (i.e., BAP+ vs. BAP-, Sasson et al., 2014; low BAP vs. high BAP, Faso et al., 2016), it is preferable to include the full range of scores in the analyses as the report may be much more informative if individuals located in the middle areas of the continuum are also taken into account (Fisher et al., 2020; Preacher et al., 2005). In this sense, Chouinard et al. (2016) found, for instance, significant correlations between AQ scores and reduced susceptibility to visual illusions. With the same data, when conducting between-group analyses ("high" and "low" groups) authors failed to arrive at the same results. On the other, the Research Domain Criteria (RDoC) initiative has proposed that categorical methods of data analysis may neglect data variability at the cost of statistical power and experimental sensitivity (Kozak & Cuthbert, 2016). In sum, these two ideas recommend that categorical analysis may result inadvisable for addressing dimensional constructs such as autism.

From a practical perspective, authors can make use of the BAP levels of expression for exploring questions in a way that might be almost impossible in clinically affected populations (Landry & Chouinard, 2016). This will be the case of neuroimaging studies which require prolonged testing sessions and multiple and high-demanding tasks to precisely delineate and track which brain areas or systems underlie specific cognitive processes. With individuals at typically functional levels even if these studies are still tedious, they are also much more feasible.

On the other hand, comorbid conditions are high in autism with estimates around the 70%, social anxiety and attention deficit/hyperactivity disorder being the most commonly reported (Simonoff et al., 2008). High comorbidity in ASD can be easily controlled thanks to the collaboration of individuals at typically functional

[5] This process consists in creating a dichotomous variable from a continuous one by splitting the scale at some point and appointed individuals above and below that point, generating two separate groups (McCallum et al., 2002).

levels since comorbidity is less frequent among them (Simonoff et al., 2008). Likewise, with samples at BAP levels, we will have the opportunity of studying autism expressions without the influences of other related disorders or associated aspects. In this line, through the BAP, researchers can easily address problems regarding the disparities in chronological and mental ages or intelligence quotient and adaptative behaviour of the participants. Many of the studies have matched clinical and control groups according to either their chronological or their mental age, even by their intelligence quotient but normally in isolation because it is hardly possible to have access to a bigger sample by paring all those aspects at the same time. However, many skills are influenced by some aspects or a conjunction of them (e.g., adults diagnosed with ASD and average-range IQ showed lower rates of independent living, Howlin and Moss, 2012; and higher rates of required support in employment and high education, Howlin et al., 2004). The use of samples at BAP levels with average-range IQ, several autistic behaviours but functionally independent, and where we normally find a correspondence between mental and chronological ages, will easily amend this issue. Finally, it will also be plausible to explore the trajectory of autism deficits over time which is one of the most important aspects of the research in ASD nowadays.

To finish this section, I would like to highlight that those studies targeting BAP severity levels complement rather than substitute research carried out with samples at ASD levels. It provides complementary perspectives that enrich the field of research, establishes foundational work, and contributes to future studies at ASD severity leve

Third Part. The Measurement of the Less Severe Expressions of Autism

Prior to conducting any study, researchers need measurement tools which enable them to characterize and select the sample according to certain targeted characteristics (i.e., autistic behaviours at BAP severity levels). This is where the need for the measurement of the BAP falls into place. However, there is still no agreed conceptualization of the BAP, and neither does a consensus in BAP measurement exist. As a matter of fact, BAP levels of expressions have been qualitatively and quantitively studied through different tools. The third part of this chapter corresponds to the revision of the measurement procedure of the BAP.

3.1 The Richness Drove to the Practice: A Historical Walk Throughout the BAP Measurement

Unlike the ASD, which counts on standardized international instruments for its assessment like the Autism Diagnostic Observation Schedule—2nd edition (ADOS-2; Lord et al., 2012) or the Autism Diagnostic Interview—Revised (ADI-R; Lord et al., 1994; Rutter et al., 2003), there are no internationally accepted instruments for assessing the BAP to date. Therefore, varied measurement tools have been applied across studies. At the end of the third part (starting on pp. 80), Table 3 resumes the main aspects of the most used measurement tools to study the BAP to date and discusses some advantages and disadvantages of their application.

In its origins, the BAP was assessed through structured and extensive personality interviews (i.e., M-PAS-R; Piven et al., 1994) to explore behavioural and genetic variability among the first-degree relatives of an individual diagnosed with ASD in an attempt to locate the underlying genes of the disorder (Bolton et al., 1994; Dawson et al., 2007; Piven et al., 1994). One of the earliest and most commonly applied tools was the Autism Family History Interview (Bolton et al., 1994), a standardized clinical interview that asked the informant detailed questions about his or her own and his or her relatives' personality and behaviours regarding the core ASD dimensions (i.e., A. Social communication deficits and B. Repetitive/rigid patterns of behaviours). However, clinical interviews implied considerable time, resulted tedious for participants as for researchers, and required qualified and trained staff. At the expense of the richness derived from interviews, their practical drawbacks for quantitative research led to the development of brief self or informant reports for measuring autistic expressions at low severity levels. An example of these screening tests was the Broader Phenotype Autism Symptom Scale (BPASS; Dawson et al., 2007) which was a combination of the parent interview and an observational rating scale that was developed for its use in genetic studies of ASD in multiplex families. Unlike the previous tools that provided a qualitative measure of the BAP, the BPASS offered a quantitative measure of autistic behaviours in four domains—social motivation, social expressiveness, conversational skills, and flexibility/range of interests—in children with high autistic levels, their parents, and siblings.

By the same time, three famous primary self-report questionnaires emerged for this purpose, the AQ (Baron-Cohen et al., 2001), the SRS-II (Constantino & Gruber, 2012), and the BAPQ (Hurley et al., 2007) which have been and still are extendedly applied. Both the AQ and the SRS are self-reported questionnaires originally developed to assess autism at clinical levels in adults with average intelligence. By contrast, the BAPQ was specifically designed to apprehend the key expressions of autism at BAP levels in adults (Hurley et al., 2007). Although many studies have found that the three questionnaires are suitable for measuring the phenotype in relatives of people with an ASD diagnosis as well as in the general population, the BAPQ has been suggested as the best instrument to identify the BAP since it was the only questionnaire specially developed for this purpose (Hurley et al., 2007; Wainer et al., 2011). Additionally, the BAPQ operationalized the BAP based on the original structure proposed by the authors of Piven et al. (1997a, b) studies. For this reason, the questionnaire is divided into three subscales referred to each of the ASD domains proposed by the DSM-IV (1994). Its items were derived from the MPAS-R and the pragmatic rating scale (PRS; Landa et al., 1992) which had been previously applied in studies exploring the BAP (Hurley et al. 2007; Piven et al., 1997a, b). Furthermore, it has been extendedly adapted to different languages and over different countries (e.g., *Canada,* Camodeca & Voelker 2016; *China*, Shi et al., 2015; *India*, Meera et al., 2015; *Israel,* Seidman et al., 2012).

3.2 Test Theory and Psychometric Premises: Measurement Problems Derived from Neglecting Validity Evidence and Further Concerns in the Psychometric Assessment of the BAP

Psychological phenomena are usually represented as latent constructs not directly observable and, so, measures developed to assess these phenomena require a process of inference validation (Flake et al., 2017). This process produces evidence that the scores of the test represent the target construct. Consequently, any measurement tool and procedure should be built upon an arduous and careful protocol of test development (Irwing et al, 2018). In the words of Flake et al. (2017, p. 371) it represents "*a process of integrating evidence to support the meaning of a number which is assumed to represent a psychological construct*". As Flake's et al. reflects, the obtention of validity evidence could be organised in three phases: substantive, structural, and external. The substantive phase is when the definition of the construct, such as the BAP, and the theory about its structure happens (Flake et al., 2017). Here, setting a strong theoretical foundation must always be at the first position. Unfortunately, problems in the psychometric assessment of the BAP start from that first stage.

As the readers may note, this dissertation has started with a thorough revision of the literature of autism in general and this includes the BAP. From that revision, readers could draw some main conclusions. One of the most important ones is that researchers have claimed that the BAP, as the phenotypic expressions of autism, should comprehend the primary dimensions of the disorder. Consequently, BAP and ASD operationalizations should be parallel and both inserted in the autism continuum at different levels of severity. In its origins, the BAP was defined by three main components (i.e., aloofness, rigidity, and pragmatic language deficits; Bailey et al., 1998) corresponding with the AS definition (i.e., detriment of social interaction, impaired communication, and repetitive and restricted behaviours and interests; APA, 1994). However, with the arrival of the DSM-V, ASD has taken a dimensional approach and its structure has been collapsed into two main dimensions (APA. 2013) while the BAP one has not been revised. In consequence, it exists a gap (mismatch) between what we understand as ASD nowadays and its phenotypical expression, the BAP. As an initial problem in construct validation, this mismatch has been hindering and influencing further evidence regarding construct validation in the field.

3.2.1 Different BAP Tests Led to Different ASD-BAP Constructs

Self-report tests have been the preferred method in personality psychology (Vazire, 2006). Their use has reduced time costs, enhanced objectivity, and allowed scale people along the autism continuums by the severity of their behaviours. However, despite having plenty of benefits, researchers and clinicians should be especially careful in the use of those tests to measure the BAP since they vary in their approach defining autism, the autistic behaviours assessed, the informant, the method of administration, and each one contemplates a specific (different) internal structure, all of which may irrevocably influence the studied outcomes. Therefore, the second concern involves problems regarding the construct and target population of each test.

Specifically, the AQ (Baron-Cohen et al., 2001) and the SRS (SRS-II; Constantino & Gruber, 2012) were originally designed for assessing autism in individuals with typical-range IQs (see Table 3). They both have developed several versions either for children, adolescents, and adults. In comparison, the BAPQ (Hurley et al., 2007), the Broader Phenotype Autism Symptom Scale (BPASS; Dawson et al., 2007), and the Subthreshold Autism Trait Questionnaire (SATQ; Kanne et al., 2012) were designed for assessing the BAP. The BAPQ originally targeted parents of children diagnosed with ASD, the BPASS unaffected siblings and parents of children diagnosed with ASD, and the SATQ general adults. All of them have been further applied to adults in the general population without a family history of autism. Other tests such as the Social Communication Questionnaire (SCQ; Berument et al., 1999; Rutter et al., 2003), the Children's Communication Checklist—2 (CCC-2; Bishop, 2003), the Childhood Asperger Syndrome Test (CAST; Scott et al., 2002), and the Social and Communication Disorders Checklist (SCDC; Skuse et al., 2005) have been applied to the exploration of autism (high-risk autism in the case of the SCQ or Asperger Syndrome in the case of the CAST) in school-aged siblings and twins of infants diagnosed with ASD.

Related to this issue, the third concern involves problems regarding the operationalization of autism inside each test as only some of them have been settled under a strong foundational theory about autism. For instance, the AQ and the SRS did not even follow the autism definition in force by the time they were created (i.e., the AS one in the DSM-IV-TR; APA, 2000) and they supported alternative theories. The AQ comprised five factors more related to cognitive processes (Social Skills, Attention Switching, Attention to Detail, Communication, and Imagination; Baron-Cohen et al., 2001). The SRS, which originally was suggested for measuring only social responsiveness, has been updated and it has included a fifth component connected with the second dimension of ASD in the DSM-V (the current five subscales are Social Awareness, Social Cognition, Social Communication, Social Motivation, and Restrictive and Repetitive Behaviors; Bruni, 2014; Constantino & Gruber, 2012). However, a unique component connected with the first dimension of the ASD is still missing and, instead of this, it includes four subscales assessing social skills. The same can be said about the CCC-2 which intended to assess pragmatic language difficulties with ten subscales (Speech, Syntax, Semantics, Coherence, Inappropriate initiation, Stereotyped Language, Use of Context, Non-verbal Communication, Social Relations, and Interests). By contrast, the SCQ, the CAST, and the SCDC included three subscales that reflect the well-known "autistic triad" and matched with former autism diagnostical definition in the DSM-IV (Social Impairments, Communication Impairments, and Repetitive or Stereotyped Behaviours).

This concern is also expanded to the tests applied for measuring BAP levels. In this regard, the BPASS comprehended four domains (Social Motivation, Social Expressiveness, Conversational Skills, and Flexibility/range of Interests) which did not match either with the original BAP operationalization (Bailey et al., 1998) or the AS (APA, 1994) or ASD (APA, 2013) ones. Something similar is observed in the SATQ (five factors: Social Interaction and Enjoyment, Oddness, Reading Facial Expressions, Expressive Language, and Rigidity). Finally, the BAPQ was developed upon the base of the BAP operationalization and, thus, it contemplated three components of impairment (Aloofness, Pragmatic Language Deficits, and Rigidity) which corresponded to autistic domains in the DSM-IV (APA, 1994). Yet, even though the BAPQ was set on a strong theory, this theory has since been outdated.

On the other hand, some authors, aware of this important problem (most of the time not the original authors of the tests), had made considerable efforts in collapsing tests' subscales and they suggested alternative models. This is the case of the AQ which alternative three-factor model has been: Social skills, Details/patterns, and Communication/mindreading (English et al., 2019; Hurst et al., 2007; Russell-Smith et al., 2011), corresponding with the previous ASD domains in the DSM-IV (1994). Nonetheless, findings regarding the alternative AQ' scores internal structure are contradictory. Moreover, even in the hypothetic case of a stable three-factor solution, it would correspond with an outdated structure of the ASD (see Table 3) which does not contemplate its dimensional nature (the same could be applied to the BAPQ as mentioned in the precedent paragraph). Finally, the SRS-II also suggests that its five subscales' scores may be grasped into two ones compatible with the DSM-V (Constantino & Gruber, 2012) to facilitate the comparison with the two dimensions of the ASD in the DSM-V: Social Communication and Interaction, and Restricted Interests and Repetitive Behaviours. Even if these comparisons bring the opportunity to determine whether an adult may meet these criteria (Constantino & Gruber, 2012), it also makes clear that the purpose of the authors of the test is to provide an ASD diagnostic questionnaire for adults. Thus, its use for assessing lower levels of severity (i.e., BAP levels) may not be adequate. About this, researchers have tended to use the BAPQ or the AQ in their studies to assess the BAP.

Nevertheless, these kinds of data-driven approaches (generally speaking, the authors look for a model with the best fit) are only an (insufficient) remedy for one of the several sources of validity evidence necessary to support the interpretation of the test scores (e.g., to provide the same number of factors involved in the ASD structure represented in the DSM-V). These internal structure adjustments do not alter the initial content of any test, which is based on the original operationalization of the construct. Consequently, if a test was originally developed upon a wrong conceptualization of the phenomena, this lack of contents will be still presented in ulterior empirical processes. To make this information more accessible for the reader, for example, if one test had not included hyper- or hyporeactivity-like behaviours of autism, changing the internal structure is not going to solve the underrepresentation of these behaviours. Consequently, the inferences made on the test scores could not be attributable to the ASD construct since the neglected contents are central autistic-like behaviours according to the DSM-V (APA, 2013). As each of these questionnaires involves a different conceptualization of the target construct, the outcomes derived from its application could drive to different inferences. Falling in this *jingle fallacies*[6] could lead to, for example, not have a complete understanding of the specific profiles and cognitive performances associated with ASD-BAP levels.

3.2.2 Measurement Problems Derived from the Mismatch between the BAP Theoretical Conceptualization and the Internal Structure of the Most Applied Tests

As tests have diverged in their operationalization of autism (either those which have targeted autism in adults with normal IQ or those which have assessed the BAP), they have not shared a common internal structure (the internal structure of each test corresponds with its subscales; test subscales have been explained in the precedent section and can be consulted in Table 3). Divergences in test scores internal structures could drive challenging interpretations of the test scores (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 2014) and could hinder further conclusions in research since each test is assessing essentially a different construct.

It seems obvious then that the internal structure should represent the operationalization of the intended construct so that researchers could draw conclusions about a construct from the test scores. However, this is not always the followed path in psychological testing. Many of the validity evidence based on the internal structure arises from questionable practices that, sometimes, are more the reflection of capitalization on chance[7] than of a theoretically driven procedure. For instance, authors have made efforts to collapse the items of the SATQ into a three-factor solution coherent with the original BAP operationalization. However, the authors took as reasonable a five-factor solution resulting from an exploratory factor analysis (see Kanne et al., 2012). Unless a test structure reflects the construct one, we cannot assure that this test is assessing what it is intended for. Furthermore,

[6] To assume that two different things are the same just because they have received the same name.

[7] Capitalization on chance occurs when authors have fitted the model to a particular sample with idiosyncratic characteristics in such a way that it may influence the particular modifications that have been performed in the model (McCallum et al., 1992).

the neglection of the theory may drive context or population-dependent validation processes (Messick, 1995). The problem with these processes lays in the replication of the findings. This could be the case of the most applied test for assessing the BAP such as the AQ, the SRS-II, or the BAPQ which have shown unstable internal structures across applications or adaptations (see Table 3).

Derived from this, most of the research in these tests has focused on providing validity evidence based on test relations with other variables but little is known about the internal structures of the tests or their content. When provided, validity evidence regarding the internal structure of the tests has shown instability between samples and studies. As displayed in Table 3, unstable test structures have been found across test applications or test versions. This could be explained partially (i) for the lack of a strong theoretical foundation of the test; (ii) because of deviations in target populations (i.e., the BAPQ originally targeted parents of children diagnosed with ASD and it has been further applied to community samples); and this has also driven to sample-dependent construct validation processes. For example, previous validity evidence on the BAPQ internal structure has pointed to dimensionality problems in one of its subscales (Pragmatic Language subscale; Sasson et al., 2013a). According to the *Standards* (AERA et al., 2014), whether two tests define the same construct differently and, thus, they contain different internal structures, those tests are assessing essentially two different things. It could also extend to structural instability. As it has been stated, "*all steps in the testing process, …, should be designed in such a manner as to minimize construct-irrelevant variance and to promote valid score interpretations for the intended uses for all examinees in the intended population*" (Standard 3.0; AERA et al, 2014, p. 63).

In an attempt to amend this problem, Wainer et al. (2011) explored the empirically-based structure for the BAP in a non-clinical sample of undergraduates from the AQ, the SRS—adult version (SRS-A), and the BAPQ using exploratory factor analyses. Based on the original conceptualization of the BAP (Piven et al., 1997a, b), they found that the scores of the thirteen factors of the three questionnaires (five from the AQ and the SRS-A, and three from the BAPQ) could be grasped into main three components: Pragmatic Language Difficulties, Aloofness, and Rigidity. These results could have indicated that part of the variance, accounted for the mentioned tests, was indeed BAP variance while, probably, there was part of these test contents and variances that could have been unrelated to the BAP. Nevertheless, this analysis should be reinterpreted upon the basis of an updated conceptualization of the BAP in line with the ASD one (APA, 2013).

3.2.3 Construct Underrepresentation and Irrelevant Content: When Tests Have Forgotten the Roots of Autism, Their Outcomes Would Blur the Construct

Finally, the fourth consideration is connected with the previous ones and resulted from the disagreement in the autistic operationalizations across tests: content concern. This last issue refers to the lack of some autistic key contents. Validity evidence based on test content takes place in the, previously mentioned, substantive phase and is normally yielded by expert judgments. Validity evidence regarding test content (Lynn, 1986; Sireci, 1998) refers to the relevance of the content of the items for assessing a specific construct and to what extent the construct has been represented adequately in the test. Ideally, a test should consider a representative sample of all the possible and relevant behaviours that conform to a psychological phenomenon.

In the specific case of autism assessment, a test should provide evidence about the relevance and the representativeness of the items for covering the whole autism continuum construct (Haynes et al., 1995; Lynn, 1986; Nunnally & Bernstein, 1994; Suen & Ary, 1989). However, studies providing this source of evidence are scarce and little is known about the importance of some test subscales which diverge from the ASD dimensions (i.e., Attention to Detail subscale from the AQ; see Table 3) for the assessment of autism.

The neglect of this validity evidence could lead to construct *misrepresentation* which explains the extent to which a test fails to capture important aspects of the target construct (Sireci, 1998). This would be the case of the AQ or the BAPQ which have not included key behaviours of the restricted and repetitive dimension (i.e., hyper- or hypo-reactivity to sensory input or unusual interest to sensory aspects of the environment). The representation of other key content may be unclear because of the way it is represented in the test structures. In this regard, the Pragmatic Language Deficits subscale of the BAPQ contemplate further language aspects rather than social communication such as restricted and idiosyncratic speech, and circumscribed topics in conversation which are now considered as key indicators of the restricted and repetitive behaviours dimension. Consequently, the content of this subscale should be reallocated inside each of the two autism severity continuums.

Finally, by including more aspects than those proposed in the manuals of reference, a test may also err on construct *irrelevance* which refers to the degree to which test scores may be affected by extraneous processes to the test's intended purposes (AERA et al., 2014; e.g., the Imagination subscale from the AQ does not directly correspond to any subdomain of the ASD definition in the DSM-V).

In my view (based on the psychometric assessment premises abovementioned), authors should make considerable efforts and supply substantial reasons for the inclusion of any subscale of any test which does not reflect directly a key autistic component according to the manuals of reference (either the DSM-V; APA, 2014; or the ICD-11; WHO, 2022) and grounded in a strong theoretical foundation. Equally, the non-inclusion of some key components should be justified upon a theoretical basis.

3.3 Aggravating the Situation: The International Adaptation of Unstable Tests

Another alarming issue is the fact that some measures have not even shared the same structure in their original version and the adapted ones (see Table 3). Consequently, even the same tests are considering the same construct differently among their versions.

In relation to what has been discussed in the previous sections, as the structures of the most commonly applied test in autism are rather a reflection of the empirical results (factors are mere groups of items that share common variance) than of the autism operationalization (or its updated dimensional one; APA, 2013), further implications have been observed. Regarding this issue, it must be noted that common factor models have been widely applied in quantitative methods in psychology. Those models are used to obtain adequate factor solutions when a test is developed. Normally, if a model does not fit adequately the first time, authors should practice *specification searches* to improve the model fit to the data (McCallum, 1986). When authors do not establish clear and theory-based strategies for model specifications in the search of the internal structure of a test, search strategies will have limitations in terms of test capability for diagnosing and correcting misspecifications (McCallum et al., 1992). In this regard, search strategies that are data-driven instead of theory-driven are inherently disposed to problems rising from capitalization on chance (McCallum et al., 1992).

Capitalization on chance hinders the generalizability and the interpretability of the results to other samples (as it happens when a BAP test is adapted to other language or sample). The problem with data-driven models (regardless of the search strategy) is that it challenges the stability of the results across applications and adaptations to other countries. According to McCallum et al. (1992), search strategies should set some priorities before conducting model redefinitions (e.g., what aspect of the model should be modified in the first place or what indicators should be included necessarily in the final model) and it could be much better whether those priorities would be established upon a strong theoretical ground.

The AERA et al. (2014) has pronounced themselves in this regard by suggesting that all test versions and adaptations to multiple languages should be thoroughly validated. The adaptation process gains meaning in this sense and the International Test Commission (ITC, 2017) has reported some guidelines for guaranteeing the quality of this process, and ulteriorly, improving the quality of the adapted versions and enabling the comparability of scores across cultures. However, despite being well-known, ITC guidelines are not usually implemented (Hernandez et al., 2020), so, even though many publishers have translated tests from one language to another, few have gathered as much data on the target language as on the original measure. Yet, to provide comparability of inferences made from test version scores lies on whether they represent the same construct as those from the original version of the test (AERA et al., 2014). For instance, whether BAPQ has been extensively applied in different countries and adapted to diverse languages (e.g., *Canada,* Camodeca & Voelker, 2016; *Chinese*, Shi et al., 2015; *Hindi*, Meera et al., 2015; *Israel,* Seidman et al., 2012; see Table 3), research has merely informed about evidence from previous studies (normally the one where the BAPQ was developed; Hurley et al., 2007) assuming that evidence is capable of being expanded to ulterior studies even in different countries (see Table 3); or reported psychometric evidence about the BAPQ's subscales (normally α was the only psychometric information reported), and validity evidence based on the BAPQ subscales relations with other variables (see Table 3). However, the validation process in every test adaptation is a necessary prerequisite for guaranteeing the interpretation of the results and it cannot be certified if no evidence is provided.

On the other hand, evidence about the differential functioning of the test versions cross-culturally should equally be reported (Hambleton et al., 2004). This extremely important evidence enables the comparability of the inferences made from test version scores. This way, authors could conclude whether a test represents the same construct in two countries and whether the psychological phenomenon is expressing itself differently. This is even more important if we consider the dimensional perspective of autism. From that perspective, we can think that is not the quality of the behaviour but its intensity that is modulated by a specific culture. In other words, the same autistic behaviours could be more or less severe according to the country and its specific culture (Berry et al., 1992; *genuine culture specifics*, Hambleton, 1994).

In this regard, a growing number of authors have highlighted the importance of assessing test functioning when applied to different languages to assure that all test adaptions preserve the equivalence of measurement among cultures (Hambleton, 1994). Even society's view of a disorder could influence the severity of the perceived symptoms, like some autistic social expressions which, for instance, appear to be less severe in Nicaragua than in Korea due to the differential impact of their particular cultural backgrounds (Kim, 2012). Consequently, cultural norms may mediate both the endorsement of symptoms and how people rate the severity of symptoms (Rogler, 1993). Coherently, two participants with the same BAP severity but from different countries or cultures could express BAP differently.

Considering this, to explore BAP levels of expression cross-culturally, at least invariant internal structure across countries should be met (i.e., configurational invariance). An invariant internal structure would prove that two subgroups of people from different cultures conceptualize the BAP in the same way (Meredith, 1993; Brown, 2015). Ideally, to further compare the obtained ratings of both populations, both test versions should be metrically invariant. This would suggest that two populations may respond to the items in similar ways (Chakraborty, 2017).

Unlike the BAPQ, autistic test stability among cultures has been studied in other questionnaires (e.g., SRS, Bölte et al., 2008; Empathy Quotient, Wakabayashi et al., 2006; see Table 3). Importantly, those tests did not show structure stability (not even configurational invariance) with their translated versions (e.g., the SRS-II has been structured in one single factor in its German version, Bölte et al., 2008 and reported five factors in the Spanish version, Constantino & Gruber, 2017; see Table 3) and this hinders the comparison of scores cross-culturally.

In conclusion, given the problems derived from the application of a test in populations different from which it originally targeted (Geisinger, 2016) as it happens with the AQ or the SRS, summed to the fact that the BAPQ is the only test that contemplates the outdated yet accepted operationalization of the BAP, and, finally, due to its extensive application and adaptation around the world, it becomes urgent that further psychometric studies tested the metric invariance of the BAPQ.

Additionally, when this dissertation started, there was not any tool that assessed the BAP in Spanish speakers. Even the BAPQ, which has been broadly applied and adapted to multiple languages, did not have a Spanish version. The Spanish adaptation of a test for measuring the BAP may open the door to multiple possibilities and would enable the expansion of BAP studies through many countries around the world. Particularly, if we consider that, in 2021, Spanish accounted for the fourth most widespread language around the world (Szmigiera, 2021) and the second most frequent native language per number of inhabitants. Consequently, both the reconceptualization of the BAP and the redefinition of a BAP measurement tool will be the aspects that my dissertation will revolve around.

Table 3

Revision of the Commonly Applied Tests for Assessing the BAP

Test	Target construct	Target population	Ground operationalization	Pros	Cons
Broader Phenotype Autism Symptom Scale (BPASS; Dawson et al., 2007)	BAP	Affected children, unaffected siblings, and parents	Four domains: ❊Social Motivation ❊Social Expressiveness ❊Conversational Skills ❊Flexibility/range of Interests	❊Provides a quantitative measure of autistic traits ❊Interview richness ❊Qualitative data ❊Structure related to the autistic dimensions in the DSM- IV (APA, 1994).	❊Long interview ❊The same interview either for children or adults as the interviewer must accommodate the patient. ❊Require trained clinicians and direct observation. ❊Not been commonly applied. ❊Does not share the two-domain structure of the ASD.
Broad Autism Questionnaire (BAPQ; Hurley et al, 2007)	BAP	Parents (originally), general adults	Three subscales: ❊Aloofness ❊Pragmatic Language Deficits ❊Rigidity	❊Structure corresponding to each of the autistic dimensions in the DSM-IV (APA, 1994). ❊Suggested instrument to identified BAP (Hurley et al., 2007; Wainer et al., 2011). ❊Good reliability and some favourable validity evidence (Hurley et al., 2007; Ingersoll et al., 2011; Sasson et al. 2013a).	❊Inadequate validity evidence based on internal structure: dimensionality problems in Pragmatic Language Deficits (Sasson et al., 2013a) and comparative fit index (CFI) indicators far from acceptable (e.g., Broderick et al., 2015). ❊Unreported factor loadings (e.g., Broderick et al., 2015; Ingersoll et al., 2011; Wainer et al., 2011).

Test	Target construct	Target population	Ground operationalization	Pros	Cons
				❊Self- and informant-report versions. ❊Applied in the general population. ❊Commonly and worldwide applied to measure BAP. ❊Many translated versions (e.g., *Chinese*, Shi et al., 2015; *Hindi*, Meera et al., 2015; *Israel,* Seidman et al., 2012). ❊Free access to the original BAPQ (Hurley et al., 2007) and some adapted versions.	❊Suggestions of items redistribution into a two-factor solution (Wainer et al., 2011). ❊Does not share the two-domain structure of the ASD.
Subthreshold Autism Trait Questionnaire (SATQ; Kanne et al., 2012)	BAP	General adults	Five factors: ❊Social Interaction & Enjoyment ❊Oddness ❊Reading Facial Expressions	❊Brief self-report questionnaire. ❊Grasps relevant expressions in broader populations. ❊Differentiate between ASD and common students.	❊Unacceptable fit for one and three-factor alternatives (Kanner et al., 2012). ❊Not very commonly applied. ❊Does not share the two-domain structure of the ASD.

Test	**Target construct**	**Target population**	**Ground operationalization**	**Pros**	**Cons**
			❉Expressive Language ❉Rigidity	❉Captures information beyond other tests (e.g., AQ; Kanne et al., 2012).	
Autism Quotient (AQ; Baron-Cohen et al., 2001)	Autism in individuals with average intelligence	Children, adolescents and adults	<u>Five factors</u>: ❉Social Skills ❉Attention Switching ❉Attention to Detail ❉Communication ❉Imagination	❉Acceptable internal consistency (Baron-Cohen et al., 2001) and strong test-retest reliability for total scores (Gravetter & Forzano, 2009). ❉AQ subscales recomposed in a three-factor model (DSM-IV; APA, 1994): Social skills, Details/patterns, and Communication/mindreading (English et al., 2019; Hurst et al., 2007; Russell-Smith et al., 2011). ❉Applied in the general population. ❉Commonly and worldwide applied to measure BAP. ❉A combination of relative and control samples suggests that AQ	❉15-20 min to complete. ❉Little knowledge of each subscale reliability (Zhang et al., 2016). ❉The only test about each subscale reliability for the original AQ counts on a very small sample (17; Baron-Cohen et al., 2001), remaining studies parts from different adapted versions. ❉An inconsistent number of factors (e.g., two-factor solutions, Hoekstra et al., 2008, 2011; Valla et al., 2010; three-factor solutions, Austin, 2005; Eriksson, 2013; Hurst et al., 2007; Ingersoll et al., 2011; Palmer et al., 2015; four-factor solutions, Saito et al., 2014; and five-factor solutions, Kloosterman et al., 2011), also through versions (e.g., two factors in AQ-Short version; Hoekstra et

Test	Target construct	Target population	Ground operationalization	Pros	Cons
				could serve to study the distribution of autistic traits (Ruzich et al., 2017). ❊Multiple test versions: self-report version, caregiver or parent-report version (for children/adolescents), short version (Hoekstra et al., 2011). ❊Translated versions (e.g., *Chinese*, Zhang et al., 2016; *French-Canadian*, Lepage et al., 2009; *German*, Freitag et al, 2007; *Italian*, Ruta et al., 2012; *Japanese*, Kurita et al., 2005).	al., 2011) and translations (e.g., three factors in the AQ-K, Freitag et al., 2007). ❊Varying scoring methods: dichotomous vs. binary scoring (Baron-Cohen et al., 2001; Ruzich et al., 2015) vs. Likert scoring, (Bishop et al., 2004; Wheelwright et al., 2010). ❊Possible gaps in some subthreshold ASD traits as originally emphasized Asperger traits (Kanne et al., 2012). ❊Does not share the two-domain structure of the ASD.
Social Responsiveness Scale—Second Edition (SRS-II (Constantino & Gruber, 2012)	Autism in individuals with average intelligence	2 years, 6 months through adulthood;	Five subscales: ❊Social Awareness ❊Social Cognition ❊Social Communication ❊Social Motivation ❊Restrictive and Repetitive Behaviors	❊Based on a sample of 1,906 individuals. ❊Total score, subscale scores, and scores for two DSM-V compatible subscales (Constantino & Gruber, 2012): Social Communication and Interaction, and Restricted	❊15-20 min to complete. ❊Qualification required. ❊Limited access (copyright by Western Psychological Services). ❊An inconsistent number of factors (e.g., one-factor solution, Constantino et al., 2000, 2004; Constantino & Todd,

Test	Target construct	Target population	Ground operationalization	Pros	Cons
				Interests and Repetitive Behaviours. ❊Facilitate the comparison of symptoms and determine whether a person may meet the most current diagnostic criteria for ASD. ❊Applied in the general population. ❊Commonly and worldwide applied to measure BAP. ❊Adequate subscales' reliability and validity with other autism subscales (ADOS, AQ, SCQ; Bölte, 2011; Bölte et al., 2008; Constantino & Todd, 2000, 2003; Virkud et al., 2008). ❊Multiple versions: Parent/teacher rating scale (from 2.5 to 18 years), self or informant-	2003), also in adapted versions (e.g., one factor in the *German* version, Bölte et al., 2008; five factors in the *Spanish* official version, Constantino & Gruber, 2017). ❊Though applied in community samples, it aims to diagnose adults with ASD (whether a person meets ASD criteria; Constantino & Gruber, 2012). ❊Does not share the two-domain structure of the ASD.

Test	Target construct	Target population	Ground operationalization	Pros	Cons
				report (adults up to 19), and short-forms. ❊Official Spanish translations of all its versions. ❊Adaptations to multiple languages (*German,* Bölte, 2011; *Vietnamese,* Nguyen et al., 2019)	
Social Communication Questionnaire (SCQ; Berument et al., 1999; Rutter et al., 2003)	Autism	High-risk children	<u>Three domains</u>: ❊Reciprocal Social Interaction ❊Communication ❊Restricted, Repetitive, and Stereotyped Patterns	❊Corresponds to the diagnosis of autism in DSM-IV (APA, 1994). ❊10 min application and 5 min evaluation. ❊Translated versions (*German,* Bölte et al., 2008, *Chinese,* Gau et al., 2011).	❊Dichotomous scoring impedes the dimensional approach and hinders the exploration of autistic traits in the general population. ❊Validation supports its use in suspected individuals rather than the general population. ❊Does not share the two-domain structure of the ASD.
Children's Communication Checklist—2 (CCC-2; Bishop, 2003)	Pragmatic Language Impairment	School-aged siblings and twins of diagnosed individuals	<u>Ten subscales</u>: ❊Speech ❊Syntax ❊Semantics ❊Interests ❊Coherence ❊Use of Context	❊Three composite scores (general communication, pragmatic, and social interaction deviance).	❊Not designed for measuring autism. ❊Does not share ASD internal structure. ❊BAP conceptualization is not clear (Ingersoll & Wainer, 2014).

Test	Target construct	Target population	Ground operationalization	Pros	Cons
			❇Inappropriate initiation ❇Stereotyped Language ❇Non-verbal Communication ❇Social Relations	❇Adapted versions (*Norwegian,* Helland et al., 2009; *Serbian*, Glumbic & Brojčin, 2012). ❇Reasonable internal consistency.	❇Only for children. ❇Inconsistencies in factor-solutions across versions (e.g., three factors in the Serbian version, Glumbic & Brojčin, 2012). ❇Does not share the two-domain structure of the ASD.
Childhood Asperger Syndrome Test (CAST; Scott et al., 2002)	Asperger Syndrome	School-aged siblings and twins of diagnosed individuals (unaffected); parent-report	<u>Three factors</u>: ❇Social Impairments ❇Communication Impairments ❇Repetitive or Stereotyped Behaviours	❇Three domains of impairment of ASC in DSM-IV-TR (Williams et al., 2005). ❇Parent or caregiver report ❇Dichotomous rating. ❇Includes special needs section.	❇Divergences in the CAST internal structure across studies in the same country (e.g., Williams, 2005, identified four factors: social behaviour and routines, speech and communication, peer relationships, and imaginative play) and adapted versions (two factors, Sun et al., 2014). ❇BAP conceptualization is not clear (Ingersoll & Wainer, 2014). ❇Targets Asperger traits. ❇Only for children.

Test	Target construct	Target population	Ground operationalization	Pros	Cons
					❊Does not share the two-domain structure of the ASD (Ingersoll & Wainer, 2014).
Social and Communication Disorders Checklist (SCDC; Skuse et al., 2005)	Autism in unaffected children	School-aged siblings and twins of diagnosed individuals	Three factors: ❊Social Impairments ❊Communication Impairments ❊Repetitive or Stereotyped Behaviours (reflecting the autistic triad) and three additional items for general behavioural problems.	❊Brief form. ❊Measures heritable characteristics. ❊Adequate internal consistency, test-retest reliability, and good discriminant validity (Skuse et al., 2005). ❊Parent report. ❊More discriminant in the general population.	❊BAP conceptualization is not clear (Ingersoll & Wainer, 2014). ❊Only for children. ❊Does not share the two-domain structure of the ASD.

Note. A revision of the main contributions of the authors in the field. For more extended results consult authors' original work.

Fourth part. A Closure of the First Chapter of the Present Dissertation

After providing a historical overview of the ASD and the BAP, I would like to close this chapter by highlighting, for the last time, the need of reconciliating the definitions of both constructs and developing a

4.1 Reconciling BAP and ASD Constructs by Providing an Updated Dimensional Operationalization of the BAP

From the view of the present dissertation, it makes no sense that the phenotypical expression of a disorder which has emerged for improving the knowledge about the aetiology of autism diverges from its founding disorder both characteristically and qualitatively. As it has been reviewed, the dimensional approach (i.e., accepting the hypothesis of the continuums of autism) offers plenty of advantages for research such as studying the traceability of ASD and how autistic-like expressions are distributed at lesser levels of severity, practical implications for the assessment and intervention on clinical levels of severity, and the empirical goodness of having access to much bigger and more heterogeneous samples. Yet, for profiting from these benefits and conducting adequate experiments, the research field needs to count on an updating operationalization of the BAP. Additionally, for drawing accurate conclusions and ensuring replicability of studies, researchers should make an effort and put the BAP under constant revision to guarantee BAP alignment with the last revised ASD definition and insert it inside the continuums of autism.

But what would a re-operationalization of the BAP aligned with the ASD one imply? In the first place, it would involve establishing a conceptual definition of the construct of interest (Delgado-Rico et al., 2012). Indeed, it has already existed a definition of the BAP, but it is outdated and, therefore, redefining it would mean to test to what extent both ASD dimensions and subdimensions as well as all specific autism behaviours disposed of in Table 1 (see, Table 1 on this chapter) could be equally applied to define the BAP. As abovementioned, some authors have already proposed that the BAP should comprise only two dimensions, namely Social and Non-Social expressions of the BAP in line with the ASD definition, where social aspects will include social impairment in social relations together with social communication, whilst non-social components constitute a rigid personality (Morrison et al., 2018; Sasson et al., 2013b).

In an attempt to simplify the research in autism in general and in BAP in particular and, also, to fix all the disagreements in how authors have approached the BAP that has been revised throughout this introduction, a new conceptualization of autism, as a severity continuum, has been proposed in this dissertation. My proposal is under the hypothesis of the autism severity continuum and the dimensional perspective of the ASD proposed in the DSM-V. Thus, the updated conceptualization of autism will revolve around only two main dimensions or domains: A. Deficits in Social Communication and Social Interaction and B. Restricted, Repetitive Patterns of Behaviours, Interests, or Activities. For each dimension, a severity continuum has been created with the autistic characteristic behaviours scaled in severity regarding how much they hinder the functionality of the person. The BAP and the ASD now take place inside each continuum and grasp specific levels of severity. Inside each continuum, we can also find, respectively, three and four sets of characteristic behaviours: Deficits in social-emotional reciprocity; Deficits in non-verbal communicative behaviours used for social interaction; Deficits in developing, maintaining, and understanding relationships (in the continuum A); Stereotyped or repetitive motor movements, use of objects, or speech; Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behaviour; Highly restricted, fixated interests; and Hyper- or hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment (in the continuum B; APA, 2013; see Table 1). I have called this operationalization "the spinal columns of autism" where only those behaviours that share relation with the rest inside a continuum and that can be scaled in severity are taken into consideration. The representation of "the spinal columns of autism" has been displayed in the 1.2.2 section, but to facilitate the reading, it can be consulted below.

With this autism reconceptualization and expansion to lesser levels of intensity, the BAP naturally grasped the same dimension as the ASD but both labels differ in the severity of expression of the autistic behaviours (corresponding the ASD with the clinical and highest levels of impairment). Regarding specific behaviours, even if some of these autistic-like expressions have already been documented in BAP (e. g., poorer interpersonal interactions, Faso et al., 2016, or strong attachment to routines, Bernier et al., 2012), others should be relocated inside the new BAP operationalization (e. g., related to echolalia, stereotypies, and hyper-sensitivity to several stimuli).

Nevertheless, the production of a BAP measurement tool requires further evidence that does not arise from the analysis of the test scores. Both the operational definition of the construct and the theoretical framework need to be scrutinized (AERA et al., 2014). The assessment of the new BAP operationalization should be performed by a group of experts in the area (see section 3.2.2; Gable & Wolf, 1993; Tilden et al., 1990) that determines how accurately the BAP has been defined by the ASD current structure and the extent to which the dimensions and subdimensions that I have delimited are relevant (Carretero-Dios & Pérez, 2007; Haladyna, 2004).

4.2 The International Measurement of the BAP

The international measurement of the dimensional structure of the BAP has highlighted some important issues. Those points deserve a throughout exploration while solutions should be attached to them. The first problem is related to the current model of autism and the theoretical substrate upon which has it been developed. As mentioned in the first part of the present chapter, autism has been originally observed and extensively reported in Western countries (Abubakar et al., 2016; de Leeuw et al., 2020; de Vries, 2016; Durkin et al., 2015; Elsabbagh et al., 2012; Hahler & Elsabbagh, 2014). The different models about autism that the DSM has proposed over the years lay their foundations in research conducted in those countries (mainly because it has been the most prolific one). Nonetheless, autistic-like behaviours have been explored worldwide (e.g., *Nigeria,* Bakare et al., 2015; *Palestina,* Basha, 2014; *Brazil,* Bordini et al., 2015) and systematic reviews have reported the international prevalence of autism from epidemiological surveys, many of them based in one of the versions of the DSM (for a review, see Elsabbagh et al., 2012). Congruently, the eventual diagnostic criteria of ASD in the DSM-V have been accepted in many countries (Amaral et al., 2011). Even when international and alternative studies are taken into consideration, the main committee of experts responsible for the neurodevelopmental disorders section in the DSM-V is still mainly formed by experts of the most prestigious universities and institutes from the United States (e.g., Johns, Hopkins University, Yale University) and the United Kingdom (e.g., Oxford University, University of Cambridge).

On the other hand, the World Health Organisation (WHO) has updated its classification of autism in the ICD-11 and is now more aligned with the DSM-V (APA, 2013). The ICD-11 was presented in the Health Assembly in May 2019 for its acceptance among Member States (WHO, 2019). In this assembly, it was advanced that the ICD-11 has defined the same two domains of ASD as the DSM-V had. However, it will not be until 1 January 2022 when modifications will come into effect. Accordingly, we cannot disregard that this is still an intermediate job, that more research should be needed and that we need to wait until 2022 to get a better consensus of the international ASD operationalization.

The second problem, which aggravates even more the situation, concerns issues I have discussed in the previous sections: the operationalization of the construct. The construct of the BAP, in particular, does not come from those models in the DSM-V or any other theoretical approach, but empirical procedures where the fit to common factor models could have made disregarding relevant BAP contents. However, as abovementioned, with data-driven procedures, authors take the risk that the operationalisation and, congruently, the measurement of the construct will be sample and/or context-dependent (the context and the sample where the test have been applied from its construction).

After all this sum of neglections, studying the stability/differential expression of the BAP across countries may seem almost impossible. The present dissertation is the reflection of how I seek this glimmer of possibility in an attempt to amend previous mistakes and bring back the bases for the correct definition and measurement of the BAP from a dimensional perspective that reconnects the phenotype with its phenomena, autism.

With this purpose, the solutions that I suggest are that both the model of classification of ASD and the dimensional model of autism (i.e., the hypothesis of the autism spectrum) that have been accepted in many countries all over the world (see two paragraphs above) should be undertaken and expanded to open the possibility to a BAP operationalization. Throughout all the pages of this first block of my dissertation, I have been providing enough support to assume that the BAP encompasses, essentially, the same expressions that the ASD does but at different levels of intensity. Equally important, the DSM-V, which has recollected extensive evidence about the disorder, is supported internationally, and I consider that there is enough evidence to be accepted as the autism model of reference. In the same line, many of the experts involved in the committee in charge of the autism diagnostic specifications in the DSM-V have conducted studies about the BAP and have expressed their opinion about the BAP represents the autistic phenotypic expressions (e.g., Cohen & Volkmar, 1997).

Understanding that ASD and BAP are both expressions of autism with an intensity nuance implies the expansion of the ASD conceptualization to the BAP. Likewise, the expansion of the ASD to lower levels involves that BAP models should be in constant revision as the ASD has been and still is. Additionally, any change in the conceptualization of ASD should be equally proved in the BAP one.

Finally, the construction of a test upon a solid theoretical foundation is naturally reached. That test should follow the recommended steps I have been talking about to provide adequate and enough evidence based on the internal structure of the test, its contents, and its relations to other variables, and lastly, guarantee the correct inferences made on its scores. As the test lays its foundations in a solid theory about autism, it should also follow a theory-driven approach and combine multiple statistical approaches to guarantee the stability of the test structure and its scores across applications and countries.

Nonetheless, prior to advance on the walk of developing a new test, by the time this dissertation started, there was not any test that measured the autistic behaviours at BAP levels in Spanish. In precedent sections, the most popular tests in the assessment of the BAP have been revised. Those tests have been worldwide applied and adapted and the research conducted with them has been very prolific. However, the lack of validity evidence needed to be solved. In this regard, the first task performed in the present dissertation will be to adapt one of the most applied tests that assess the BAP to Spanish. This will be made upon the basis that it is fundamental to translate the most popular tests in the assessment of the BAP to the Spanish language (starting by one of them) and explore whether the Spanish operationalization of the BAP corresponds to the English one (the original operationalization of the BAP) and to provide further validity evidence. In this sense, all the preceding reviewed tests have been developed in English and, later on, worldwide adapted (e.g., the Chinese version of the BAPQ, Shi et al., 2016, the Hindi version of the BAPQ, Meera et al., 2015; see Table 3, pp. 80) and, besides the problems regarding validity evidence in the original and the translated versions, adaptations have not followed the recommended ITC recommended procedures. Taking into consideration these possible inconveniences, in the case the adaptation of the test will show adverse evidence, I will proceed to search which contents of the most applied tests in the BAP could suit the updated BAP operationalization. Regarding this, whether the items of the most applied tests would adjust to the new BAP dimensional operationalization, the process of constructing a new test will be eased.

Finally, if the worst scenario is reached, this dissertation will try to end by presenting a new international test as the main outcome. Nevertheless, in this case, the construction of this measurement tool will be based on an internationally accepted ASD/BAP operationalization, provided in the DSM-V, and developed simultaneously in two countries, Spain and the United Kingdom. These two, *a priori*, trivial features could be the base to expand the international research on ASD and BAP since this measurement tool will allow reconnecting both concepts. Furthermore, the Spanish and English versions will ease the adaptation of the test to all the countries where Spanish and English are the official languages.

CHAPTER TWO. GENERAL AND SPECIFIC OBJECTIVES OF THE PRESENT DISSERTATION

General Objective

The implementation of the present dissertation was primarily motivated by the necessity of having access to a measurement tool of the BAP in the Spanish language. By the time that this dissertation started, there was not any available test that assess the BAP in Spanish speakers. In accordance with all the information provided in Chapter One, a test that assesses the BAP should count on adequate psychometric properties and validity evidence. Ideally, it should correspond to the newest and most updated ASD definition in the DSM-V as the BAP represents the endophenotype of this disorder and it should maintain those psychometric properties of the original version in its Spanish one.

Specific Objectives Presented into Studies

For achieving this general purpose, I will establish series of specific objectives which will be presented hereunder ordered by empirical studies. It should be noted that studies are linked in a way that the results of the first study give cause for the second one and so on.

Study 1. "Psychometric Properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements"

In my first study I established three main objectives:

I. To adapt the BAPQ, as the original test for assessing the BAP, to the Spanish language (the BAPQ-SP). This test has been translated to several languages but there was not any version of that test or any other for Spanish speakers at the moment that this dissertation started.

II. To analyse the psychometric properties as well provide evidence regarding the internal structure of the BAPQ-SP.

III. Finally, to establish the relations of the BAPQ-SP factors with other variables to gather validity evidence regarding the factor relations with other variables.

Study 2. "Applying a Rasch Model to the BAPQ: Item Severity Analysis and Test Differential Functioning of the English and Spanish Versions"

Study 1 disclosed some problems in the internal structure of the BAPQ-SP scores mainly caused by the Pragmatic Language subscale while the Aloof and Rigid ones showed unidimensionality. Following the dimensional approach of autism that I take in my dissertation, I consider it relevant to study the remaining two subscales employing a measurement model more suitable for the dimensional operationalization of the BAP. Thus, it can be tested 1) whether the items of Aloof and Rigid subscales and the persons measured with them could conform to severity hierarchies and 2) if the item severity levels match with those expressed by the general population (low BAP levels; see Figure 3). My specific objectives in this study are:

I. To analyse the psychometric properties of both the Spanish and the English versions of the BAPQ in light of the Item Response Theory by applying Rasch models. This approach lets us know about the scalability of the items, their severity in both versions as well as its adequacy for the targeted sample among other advantages.
II. To study, for the first time, the differential test and item functioning between both versions in terms of severity (how much the severity hierarchies change between versions and whether the specific items are located in a similar point of the continuum of severity).

Study 3. "Is it possible to assess the new two-domain definition of the Broad Autism Phenotype using the available measurement tools?"

Together, results from Studies 1 and 2 showed adverse evidence based on the internal structure of the scores of both BAPQ versions (the BAPQ-EN or original BAPQ, and the BAPQ-SP or Spanish version). In this Study 3, my main purpose is to accommodate the items of the three most applied tests to assess the BAP worldwide into its new operationalization. Congruently, the specific objectives are:

0. Baseline. To provide theoretical evidence about the new operationalization of the BAP (Ingersoll et al., 2011). As a previous stage, a group of experts revise and rate the operationalization of the BAP proposed in Study 3.
I. To provide empirical evidence regarding the empirical two-domains structure of the BAP collapsing scales' scores of the three most applied tests for studying the BAP.

II. To provide theoretical evidence whether the items of these three tests could be reshaped in a two-factor structure congruent with the newest BAP operationalization according to the ASD definition in the DSM-V. Items are reassigned to one of the seven BAP subdomains included in its new operationalization and, then, a broad group of autism experts evaluates different aspects of the items. This objective is also divided into two subsequent tasks:
   a. The itemmetric properties as the formal aspects of the items are revised and only items with adequate formal properties are maintained in the convergent pool (Angleitner et al., 1986).
   b. Then, the content of the items is analysed with a group of autism experts who rate the items according to their relevance for measuring the BAP.

III. To rate the representativeness of the BAP structure once the final pool of items has been obtained.

Study 4. "The Broad Autism Phenotype - International Test (BAP-IT): a two-domain based test for the assessment of the broad autism phenotype"

Since Study 3 resulted in a very small pool of items and experts pointed to the underrepresentation of some of the core BAP subdomains, in Study 4, I aim to design, develop, and pilot a test, the Broad Autism Phenotype—International Test (BAP-IT) in two countries simultaneously, Spain and the United Kingdom. The BAP-IT will target the updating BAP operationalization proposed in Chapter One and it will be a test for the measurement of a two-domain structure of the BAP paralleling the current ASD one. The specific objectives are:

I. To obtain the BAP-IT simplest factor structure (two correlated dimensions with only primary loadings; Brown, 2015) with an adequate content representation of the seven ASD subdomains (see Table 1; Chapter One).

II. To test whether that simplest factor structure results invariant across groups (Spanish and English groups). Specifically, both configural and metric invariances between the Spanish and British versions of the BAP-IT are aimed to be met.

III. To study further psychometric properties like reliability and validity evidence concerning other variables.

BLOCK TWO:

EMPIRICAL STUDIES

CHAPTER THREE. STUDY 1:

PSYCHOMETRIC PROPERTIES OF THE SPANISH VERSION OF THE BROAD AUTISM PHENOTYPE QUESTIONNAIRE: STRENGTHS, WEAKNESSES, AND FUTURE IMPROVEMENTS

Manuscript published as:

Godoy-Giménez, M., González-Rodríguez, A., Cañadas, F., Estévez, A. F., & Sayans-Jiménez, P. (2018). Psychometric Properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements. *Journal of autism and developmental disorders*, *48*(3), 770-783. https://doi.org/10.1007/s10803-017-3438-0

Psychometric Properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements

Marta Godoy-Giménez[1], Antonio González-Rodríguez[1], Fernando Cañadas[1,2], Angeles F. Estévez[1,2], and Pablo Sayans-Jiménez[1]

[1] Department of Psychology, University of Almeria, Spain

[2] CERNEP Research Center, University of Almeria, Spain

Author Note

Correspondence should be addressed to Pablo Sayans-Jiménez, Department of Psychology, University of Almeria, 04120 Almeria, Spain, Phone number +34 950 015947, E-mail: psj242@ual.es or Angeles F. Estévez, Department of Psychology, University of Almeria, 04120 Almeria, Spain, Phone number +34 950 214626, E-mail: mafernan@ual.es

Abstract

The Broad autism phenotype (BAP) refers to a set of subclinical behavioural characteristics qualitatively similar to those presented in autism spectrum disorders (ASDs). The BAP questionnaire (BAPQ) has been widely used to assess the BAP both in relatives of ASD people and within the general population. The current study presents the first Spanish version of the BAPQ (BAPQ-SP) and analyses its psychometric properties, including validity evidences based on the BAPQ scores relationship with other variables. Our results only support the use of the Aloof and Rigid sub-scales to assess this phenotype, whereas Pragmatic Language sub-scale seems to be the main source of misfit. This research represents a first step in the study of the BAP features in the Spanish population.

Keywords: broad autism phenotype (BAP), Spanish version of the Broad Autism Phenotype Questionnaire (BAPQ-SP), confirmatory factor analysis (CFA), validity evidence.

Psychometric properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by atypical and persistent impairment in social interaction and communication, as well as a pattern of repetitive, restricted behaviours and interests (American Psychiatric Association [APA], 2013). ASD is a highly heterogeneous set of disorders with wide variations in symptom severity, intellectual level, and functional disability (Geschwind, 2009), caused by multiple possible etiological underpinnings, which remains poorly understood (Betancur, 2011; Bruining et al., 2010; Jeste and Geschwind, 2014; Lenroot and Yeung, 2013). As a consequence, the importance of accurate diagnoses has never been greater, particularly given also the growing prevalence (Fombonne, 2009), and the considerable family and societal costs (Ganz, 2007). Furthermore, undertaking the autism as a spectrum disorder remarks that the autistic traits could extend to levels of non- clinical significance within the general population (Landry and Chouinard, 2016). In fact, if we understood autistic pat- terns as dimensional diagnostic criteria, they would obviously vary from subclinical to clinical levels of expression, throughout the general population (Constantino and Todd, 2003; Ronald et al., 2006; Ruzich et al., 2015).

Some studies have employed the concept of broad autism phenotype (BAP) to refer to a set of subclinical characteristics qualitatively similar to those presented in ASD, but milder in their expression. The BAP was first observed by Leo Kanner and Hans Asperger in 1940s; these authors reported that parents showed behavioural features similar to those typical in their autistic offspring. More specifically, Kanner (1943) observed subtle manifestations of the autistic traits in some of his patient's parents such as an obsession with details, social awkwardness, and rigid behaviours. Later on, several studies have revealed social-communication, personality, and cognitive characteristics similar to those presented in ASD in parents, siblings, and other extended relatives of people with an ASD diagnosis (i.e., enhanced aloof, rigid, and hypersensitive personality traits; increased rates of internalizing psychiatric conditions; and, deficits in pragmatic language skills as well as in social reciprocity; for reviews, see Bailey et al., 1998; Pisula and Ziegart-Sadowska, 2015; Sucksmith et al., 2011).

Most of the research on the BAP has deeply focused on genetic liability for autism by examining autistic traits in a subgroup of first- and second-degree relatives of ASD individuals (for a review, see Bailey et al., 1998). Recently, some authors have suggested that since ASD is diagnosed based on behavioural methods, the study of those people who exhibit subclinical autistic traits could be a useful tool for understanding the relationship between behaviours and traits which are similar to those of ASD (D'Cruz et al., 2013; Landry and Chouinard, 2016). Thus, behavioural, neuropsychological, and personality profiles have been explored in parents and siblings of ASD patients (Klusek et al., 2014; Losh et al., 2008, 2009; Ruzich et al., 2015).

However, innovative studies have highlighted that the BAP also exists within the general population (Baron-Cohen et al., 2001; Constantino and Todd, 2003, 2005), suggesting its relevance as a research target. Actually, recent studies have already provided meaningful information for a range of characteristics related to autism which are continuously distributed within the general population (Wainer et al., 2011), including restricted and repetitive interests and hobbies (Baron-Cohen et al., 2001), atypical visuospatial and cognitive performance (Grinter et al., 2009; Stewart and Austin, 2009), reduced gaze reciprocity (Chen and Yoon, 2011), and abnormal speech perception (Stewart and Ota, 2008) as well as recognition from faces and body language (Ingersoll, 2010). It is important to note that the studies focused on this population are still scarce but highly needed given that their results could help to gain a better understanding of the developmental trajectories of autistic tendencies, as it has been suggested in Landry and Chouinard (2016) research, which confirmed the presence of traits associated with ASD in the typically developing population. But before conducting these studies, it is necessary first to develop specific tools for measuring the BAP in the general population.

Measuring BAP

Initially, the BAP was assessed through clinical interviews with the purpose of studying behavioural and genetic variability among the first-degree family members of an individual with ASD in an attempt to locate the underlying genes of the disorder (Bolton et al., 1994; Dawson et al., 2007; Piven et al., 1994). Later on, three primary self-report questionnaires have emerged for measuring BAP characteristics—the Autism Quotient (AQ), the Social Responsiveness Scale (SRS), and the Broad Autism Phenotype Questionnaire (BAPQ)—none of them have been translated or validated in Spanish-speaking populations. The AQ is a self-report questionnaire originally developed to assess ASD in adults with average intelligence and it is comprised of five sub- scales: social skills, attention switching, attention to detail, communication, and imagination (Baron-Cohen et al. 2001; Ingersoll et al., 2011). Likewise, the SRS is a questionnaire which assesses ASD as a single trait—social reciprocity— that is made up of five subscales: social awareness, social cognition, social communication, social motivation, and repetitive and restrictive behaviour (Constantino and Gruber, 2007). Finally, the BAPQ was designed to apprehend the key features of the BAP in adults comprising three dimensions: Aloof, Rigid, and Pragmatic Language (Hurley et al., 2007). Although many studies have found that the AQ, the SRS, and the BAPQ are suitable for measuring this phenotype in relatives of people with an ASD diagnosis, the BAPQ has been suggested as the best instrument to identify BAP individuals since it was the only questionnaire specifically developed for this purpose (Hurley et al., 2007; Wainer et al., 2011). Its items were derived from the modified personality assessment schedule—revised (MPAS-R) and the pragmatic rating scale (PRS) which had previously been used in studies exploring the BAP (Hurley et al., 2007; Piven, Palmer, Jacobi et al., 1997; Piven, Palmer, Landa et al., 1997). The questionnaire is divided into three subscales referred to each of the three dimensions. According to Hurley et al. (2007, p. 1681) aloof personality refers to a limited interest in or enjoyment of social interaction; rigid personality relates to resistance, and/or difficulty adapting to change; and, pragmatic language problems comprise deficits in the social use of language that lead to difficulties with effective communication and/or conversational reciprocity.

Overall, the BAPQ scores have shown good reliability and validity (see Hurley et al., 2007; Ingersoll et al., 2011; Sasson, Lam, et al., 2013). Several studies have provided validity evidence supporting the inferences made on BAPQ scores mainly in terms of differential scores among groups (e.g., Hurley et al., 2007; Meera et al., 2015; Sasson, Lam et al., 2013; Shi et al., 2015). BAPQ scores have also supportive validity evidence based on their relations with other variables. In fact, this questionnaire is well-correlated with the AQ and the SRS (Ingersoll et al., 2011; Sasson, Nowlin et al. 2013), and additional criterion evidence have been found both in a non-clinical sample of college student (Ingersoll et al., 2011; Wainer et al., 2011) and within the general population (Sasson, Nowlin et al., 2013). This evidence showed, for example, that subthreshold autism-social characteristics within the general population, measured by using the BAPQ, were connected to poorer social skill and social- cognitive ability (Sasson, Nowlin et al. 2013). Likewise, these characteristics were related to a poor performance on an aggregate of social-cognitive aspects such as face recognition, affect processing, and theory of mind abilities.

However, not all the validity evidence clearly supports the inferences made on BAPQ scores. Validity evidence based on the internal structure is probably one of the BAPQ features that needs more attention in future research. When factor analysis is performed the main goal is to achieve a simple structure (i.e., factors should be defined by indicators loading highly on them and cross-loadings or factor overlapping should be avoided; see Thurstone, 1947) that enhances the interpretation of the scores, even more, when coarse factor scores (averaging or summing the unweighted raw of indicators with salient loadings on each factor) are used to determine Aloof, Rigid, and Pragmatic Language. Nevertheless, BAPQ factor structure has shown comparative fit index (CFI) indicators far from acceptable limits (e.g., Broderick et al., 2015). Results have also shown no evidence of a simple factor structure, sometimes because there are several items with cross-loadings or without loadings at least in one factor (e.g., Sasson, Lam et al. 2013) and, other times, because factor loadings are not reported (e.g., Broderick et al., 2015; Ingersoll et al., 2011; Wainer et al., 2011).

Regardless of these internal structure problems, the BAPQ appears to be the best questionnaire to explore the BAP in the general population, as mentioned above (Ingersoll et al., 2011). For this reason and given that none of the three questionnaires is available for Spanish speakers, the purpose of this study is to present a new Spanish self-report version of the BAPQ. To do this, we have translated and adapted the BAPQ to Spanish speakers and its psychometric properties have been tested in a non-clinical sample of college students. In this study, we specifically emphasize the information related to the internal structure. If the reported lack of fit related to the internal structure is confirmed here, we will provide information regarding its causes and possible solutions with the aim of finding a measurement tool with empirically supported inferences based on the BAPQ specific dimensions.

In the present study we will perform confirmatory factor analysis (CFA) to test BAPQ simplest factor structure (three factors with only primary loadings). We will also analyse and describe global and local fit, with special attention to modification indices in order to provide future directions to improve this questionnaire. It is important to note that Sasson, Lam et al. (2013) found that when the self-report version of the BAPQ was applied in a sample of parents of people diagnosed with ASD, five out of seven salient cross-loadings were caused by the items of the Pragmatic Language dimension (BAPQ items 2, 7 11, 29, and 34). Based on this study, our main hypothesis is that one of the possible reasons underlying the lack of fit of BAPQ internal structure might be due to the malfunctioning of Pragmatic Language items.

On the other hand, once we have identified an acceptable simple factor structure, validity evidence based on the BAPQ scores relationship with other variables will be tested. Specifically, the NEO Five-Factor Inventory (NEO-FFI; Costa and McCrae, 1992) dimensions of extroversion, openness, and agreeableness and the three subscales of the Pragmatic Awareness Questionnaire (PAQ; Rodríguez-Muñoz, 2012) will be regressed on BAPQ dimensions. NEO-FFI dimensions (NEO Five-Factor Inventory is the short version of the NEO-PI-R) were selected based on previous studies with parents with ASD children that found a negative correlation between Aloof (or social features measured by the MPAS or the MPAS-R) and the Extraversion factor measured by the NEO-PI-R, and between Rigidity (or rigid personality) and the Openness and Agreeableness dimensions (Costa and McCrae, 1995, as cited in Amaral et al. 2011, p. 463; Piven, Palmer, Jacobi et al., 1997; Piven, Palmer, Landa et al., 1997). Later on, two studies exploring the relationship between autistic traits and personality in university students by using the AQ and the Personality Mini-Markers (Austin, 2005) or the NEO-PI-R (Wakabayashi et al., 2006), have also found that AQ score was negatively correlated with Extraversion and Agreeableness. Taking into account these findings as well as the definition of the three dimensions assessed by the BAPQ (Aloof personality, Rigid personality and

Pragmatic language problems; Hurley et al., 2007, p. 1681) we hypothesize that Aloof personality will have a high negative relationship with Extraversion and a moderate negative relationship with Agreeableness, whereas Rigid personality will be weakly negatively related to Openness. The PAQ (Rodríguez-Muñoz, 2012) will be used to test the convergent validity of the Pragmatic Language scores. In this case, moderate/high negative relationships among the PAQ subscales and the Pragmatic Language dimension are expected whereas Aloof and Rigid dimensions should show low negative relations with the PAQ subscales.

Method

Participants

Three hundred and forty-nine undergraduates participated in this study. Gender composition was 26.1% of men, ranging in age from 18 to 52 years, M (SD) = 21.56 (4.63), and 73.9% of women ranging in age from 18 to 47 years, M (SD) = 20.88 (4.01). Written informed consent was obtained from all participants. They all were informed with respect to the purposes of the study and none of them had previous experience with any of the questionnaires applied (BAPQ, PAQ, and NEO-FFI). Incidental and snowball sampling was performed. The inclusion criteria were to have no history of autism or any related developmental disorders in first-degree relatives. Those participants who did not meet the inclusion criteria, have severe medical or genetic conditions, or history of psychiatric disorders (e.g., depression, schizophrenia…) were excluded. This study received ethics approval by the local Human Research Committee and was conducted in accordance with the approved guidelines and the Declaration of Helsinki.

With the 12 items per BAPQ factor this sample size meets the requirements suggested by MacCallum et al. (1999) to obtain stable factor solutions. Additionally, a priori power analysis was performed to determine what would be the minimum sample size to detect very small effect sizes ($R^2 = .05$) in a linear multiple regression with three predictors using Gpower software (Faul et al., 2009). With an alpha = .05 and power = .80 the required sample size to detect the mentioned effect size was n = 253. Nevertheless, we recruited a bigger sample because this study is framed in a larger research.

Instruments/Measures

Broad Autism Phenotype Questionnaire, Spanish Version (BAPQ-SP)

The BAPQ (Hurley et al., 2007) is a screening instrument for the assessment of the personality and language characteristics identified as encompassing the BAP in parents of individuals with ASD (Ingersoll et al., 2011). The BAPQ includes 36 items grouped into three subscales (Aloof, Rigid, and Pragmatic Language) and offers both a score for each subscale and an overall score, which reflects the average of all items (Hurley et al., 2007). Each of the subscales is comprised of 12 of the total 36 items and it is related to each of the BAP traits proposed by Hurley et al.: aloof personality, rigid personality, and pragmatic language difficulties. The BAPQ items demand participants to rate how frequently each statement (e.g., “I like being around other people”) applies to them on a scale from one (statement applies very rarely) to six (statement applies very often). This scale (1–6) does not offer neutral option as a response (Hurley et al., 2007). Fifteen of the total 36 items become reverse scored items (1, 3, 7, 9, 12, 15, 16, 19, 21, 23, 25, 28, 30, 34, and 36). There are two versions of the BAPQ (Hurley et al., 2007); in this study we only used the self-report version.

BAPQ adaptation to Spanish language was performed following the guidelines proposed by Muñiz et al. (2013) and the International Test Commission Guidelines on Adapting a Test (http://www.intestcom.org). The English version of the BAPQ was translated into Spanish by an official translator. Simultaneously, the BAPQ was also translated into Spanish by one of the researchers of the present study. Then, the concordance of each item from the two translations was ana- lysed. In this comparison, we discussed the equivalence of the cultural aspects mentioned in the BAPQ to the Spanish population, their grammatical concordance, and the quality of the wording of the items of the questionnaire. Subsequently, a second researcher translated the Spanish version back to English. Both researchers involved in the translation of the BAPQ were Spanish native speakers. After comparing the back-translation version to the original text and dis- cussing discrepancies in the retranslation with the BAPQ English version, a final version was established. As soon as the final version was obtained, a mixed panel conformed by three experts (one for each of the domains in the questionnaire—namely, Aloof personality, Rigid personality, and Pragmatic Language—being one of them also an expert in psychometrics applied to transcultural contexts, and trained in Educational settings), valued the adjustment of each of the items to their corresponding domains and the sense of the relation, that is, if they have a positive or a negative relation to the domain. The administered BAPQ-SP can be seen in Appendix 1 (Appendix Section).

Pragmatic Awareness Questionnaire (PAQ; Rodríguez-Muñoz, 2012)

This questionnaire is a self-report measure of pragmatic language in Spanish population. It is composed of 26 items that are answered using ordinal scores following a 5-point Likert scale where: 1 is very bad, 2 is bad, 3 is regular, 4 is good and 5 is very good. For its elaboration, the PAQ (Rodríguez-Muñoz, 2012) was based mainly on the Prutting and Kittchner (1987) pragmatic protocol. PAQ items are organized into three pragmatic factors: Enunciative (Cronbach α in this study .64), Textual (Cronbach α in this study .81), and Interactive (Cronbach α in this study .88); a score is given for each factor and an overall composite score can also be obtained. None of the items requires reversed scoring. The higher the score of each of the three factors, the better the pragmatic language is.

The NEO Five—Factor Inventory (NEO-FFI, Spanish Version; Costa and McCrae, 1999)

This inventory is a short well-established and widely used measure of the big five model of personality proposed by Costa and McCrae (1989). It includes 60 items that yield an overall score composed of five scales: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Each of the scales is comprised of 12 of the total 60 items (24 are reverse scored). Participants have to rate each item or statement according to their character on a nominal scale from A (totally disagree) to E (totally agree). Only the Extraversion (Cronbach α in this study .90), Agreeableness (Cronbach α in this study .74), and Openness (Cronbach α in this study .77) dimensions were considered in the present study due to their theoretical relation to the aloof and rigid BAPQ factors. Higher scores imply higher levels of Extraversion, Agreeableness, and Openness.

Procedure

Trained staff administered, in groups of 30 participants, a booklet containing a set of six self-report questionnaires: the three just described in the instruments section (BAPQ, PAQ, and NEO-FFI), and three additional questionnaires out of the scope of this study. Participants answered the survey in a quiet room free from distraction. There was no time limit (the approximate time necessary to complete the questionnaires was 40–50 min) and each participant received course credits in return. The trained staff confirmed that all the participants were participating voluntarily and informed them that their answers would be handled only for scientific purposes. Participants were also notified they could stop their participation at any time and that their anonymity and confidentiality was guaranteed. Furthermore, all of them signed an informed consent.

Two booklet versions (model A and model B) were made by changing the order of two of the three additional questionnaires although the BAPQ was always the first in filling and the PAQ always preceded the NEO-FFI. Participants received randomly one of the two models. First, participants read a letter from the researchers endorsing the aims of the research and the informed consent. Following signing the consent, the instructions to complete the whole survey were explained. No questions were allowed to be asked to the researchers. After finishing the set of questionnaires, participants were asked to check that they had completed all the questions and that they had answered them appropriately and honestly. Finally, some socio-demographic variables (name, age, gender and educational level—Degree, Master degree or PhD) were registered.

Data Analysis

The descriptive statistics of all BAPQ indicators as well as the descriptive statistics of PAQ and NEO-FFI total scores were analysed. Outlier detection was done using Mahalanobis distances. According to Hair et al. (2009), an observation can be considered as an outlier if the coefficient D2/df exceed the value of 3 or 4 in big samples. CFA and structural equations models were performed to test the BAPQ internal structure and its relationships with PAQ and NEO-FFI scales. Parameter estimations were performed using the maximum likelihood. Missing data was treated with Full Information Maximum Likelihood. We assigned the latent factors metric by fixing the mean to .00 and variance at 1.00 for all the latent variables. All the correlations between factors were freed and all the cross-loadings were fixed to 0.

Fit to the models was checked using the Chi square test, the Comparative Fit Index (CFI), and the Root Mean Square Error Approximation (RMSEA; the confidence interval was set at 90%). We consider good fit when RMSEA ≤ .05 or CFI ≥ .97 (Schermelleh-Engel et al., 2003); adequate fit when RMSEA is close to .06 or CFI ≥ .95 (Hu and Bentler, 1999). RMSEA values between .08 and .10 and CFI values between .95 and .90 can be considered as an acceptable fit. When RMSEA is higher than or equal to .10 and CFI values are lower than .90 the model should be discarded (Brown, 2006). The 95% confidence intervals were estimated using MPlus (based on the application of the delta method; Raykov and Marcoulides, 2004). Analyses were performed using SPSS v19.0 and MPlus v7.0.

Validity Evidence Based on the Internal Structure

We tested a first order factor model with three dimensions, each one composed by the Aloof, Rigid, and Pragmatic Language items respectively (BAPQ-Model; the data set contained 315 complete cases). Omega coefficients were estimated to check the score reliability on each dimension (McDonald, 1999). Adequate omega values should be higher than .80 (Raykov and Marcoulides, 2011). Additionally, due to the misfit of the tested model and based on Sasson, Lam et al.'s (2013) results, we decided to test a new first order factor model including only the Aloof and Rigid dimensions (AlRig-Model; the data set contained 329 complete cases). With this model we try to test whether the only psychometric problem of the BAPQ questionnaire are the Pragmatic Language items and dimension or, on the contrary, Aloof and Rigid items and dimensions also show severe malfunctioning. Modification indices higher that ten were analysed in both models in order to provide general guidelines to improve BAPQ internal structure.

Validity Evidence Based on the Relationships with Measures of Other Variables

Two structural models were created to test the relationships between BAPQ dimensions and PAQ and NEO-FFI scales. Extraversion, Agreeableness, Openness, Enunciative, Textual, and Interactive dimensions were regressed on BAPQ factors of BAPQ-Model and AlRig-Model (BAPQ-StrModel and AlRig-StrModel models respectively). The dataset in the BAPQ-StrModel contained 299 complete cases whereas in AlRig-StrModel the data set contained 311 complete cases. In order to discard any potential effect of the gender variable we have included measurement and structural invariance tests among genders for each measurement and structural model. Gender invariance will be assumed if CFI differences between models are lower than .01 (Cheung and Rensvold, 2002).

Results

The descriptive statistics for all the variables can be seen in Table 1. The sample means of the present study were lower than those observed in the PAQ original sample, and similar to those found in the NEO-FFI. No outliers were identified according to Hair et al.'s proposed cut- off (2009). No variables showed extreme kurtosis or skewness.

Validity Evidence Based on the Internal Structure

Fit indicators for BAPQ-Model were: $\chi 2$ (591) = 1460.22, $p < .001$, the RMSEA was .07 [.06, .07] and the CFI was .74. This measurement model showed extremely low CFI values, suggesting that it should be discarded. Measurement invariance among genders was confirmed for this model, ΔCFI = .002. Factor loadings and modification indices can be seen in Tables 2 and 3. Factor loadings in absolute values were moderate-high M = .62, range = [.31, .80] for Aloof dimension, moderate M = .50, range = [.29, .72] for the Rigid dimension, and moderate-low M = .31, range = [.03, .73] for the Pragmatic Language dimension. Four out of 12 items of Pragmatic Language showed non-significant loadings. Omega reliability estimations [95% CI] were .89 [.87, .90] for Aloof factor, .79 [.77, .82] for Rigid factor, and .54 [.48, .59] for Pragmatic Language. Cronbach's alfa reliability estimations were .85 for Aloof factor, .79 for Rigid factor, and .62 for Aloof factor. All correlations among factors were statistically significant ($p < .001$): − .44 for Aloof-Rigid, −.80 for Aloof-P. Language, and .47 for Rigid-P. Language. Given that the Aloof factor includes mostly reverse scored items and that none of the items were inverted in this study, the direction of the Aloof relationships is reversed (i.e., a high score in this dimension indicates lack of aloofness or high interest in or enjoyment of social interaction; this is applicable henceforth). Subscale intercorrelations were comparable to the BAPQ original intercorrelations found in the control group (Hurley et al., 2007) although Aloof-P. Language correlation was bigger than expected. All these results make inadvisable to use Pragmatic Language observed scores to represent or infer this construct.

Additionally, when local fit related to factor loadings was explored, six out of 12 items of Pragmatic Language showed modification indices higher than 10, whereas only two items of Aloof and Rigid, respectively, showed local misfit. Finally, modification indices related to potential correlations between item uniqueness pointed again to Pragmatic Language items as the main contributors to the global misfit of the BAPQ measurement model (i.e., BAPQ-Model).

Table 1

Descriptive Statistics for all the BAPQ Items and Related Variables

	M(SD)	SK	K		M(SD)	SK	K
Item 1	5.01 (0.99)	-0.96	0.58	Item 22	2.40 (1.15)	1.08	1.18
Item 2	2.73 (1.26)	0.75	-0.02	Item 23	2.13 (1.00)	0.84	0.62
Item 3	3.05 (1.19)	0.40	-0.28	Item 24	2.44 (1.15)	0.74	0.19
Item 4	2.00 (1.21)	1.58	2.18	Item 25	4.42 (1.21)	-0.47	-0.53
Item 5	2.21 (1.17)	1.09	0.92	Item 26	2.27 (1.27)	1.19	1.07
Item 6	2.43 (1.16)	0.80	0.44	Item 27	1.82 (1.05)	1.40	1.71
Item 7	4.38 (1.07)	-0.40	-0.11	Item 28	4.70 (1.16)	-0.63	-0.27
Item 8	2.04 (1.19)	1.25	1.10	Item 29	2.02 (0.94)	1.35	3.04
Item 9	4.82 (1.13)	-0.94	0.68	Item 30	3.10 (1.09)	0.48	0.20
Item 10	2.25 (1.22)	1.08	0.78	Item 31	2.28 (1.17)	0.83	0.46
Item 11	2.08 (1.02)	1.03	1.20	Item 32	2.96 (1.35)	0.43	-0.57
Item 12	4.13 (1.29)	-0.42	-0.62	Item 33	3.38 (1.31)	0.32	-0.53
Item 13	2.08 (1.01)	0.95	0.78	Item 34	4.16 (1.05)	-0.25	-0.45
Item 14	2.33 (1.13)	1.15	1.42	Item 35	2.97 (1.14)	0.58	0.13
Item 15	4.02 (1.09)	-0.02	-0.73	Item 36	4.81 (1.13)	-0.81	0.17
Item 16	4.21 (1.34)	-0.34	-0.74	Extraversion	3.74 (0.72)	-0.66	0.01
Item 17	2.75 (1.36)	0.79	-0.07	Agreeableness	3.53 (0.52)	-0.43	0.22
Item 18	2.26 (0.99)	0.79	0.75	Openness	3.64 (0.59)	-0.15	-0.52
Item 19	4.88 (1.09)	-0.65	-0.45	Enunciative P.	4.05 (0.58)	-0.39	-0.17
Item 20	3.23 (1.46)	0.25	-0.84	Textual P.	4.02 (0.60)	-0.40	-0.21
Item 21	4.54 (1.10)	-0.56	0.10	Interactive P.	3.95 (0.51)	-0.14	0.14

Note. M = Mean; SD = Standard Deviation; Sk = Skewness; K = Kurtosis.

Accordingly, we ran an alternative measurement model in which Pragmatic Language items were excluded. Fit indicators for AlRig-Model were: χ2 (591) = 579.78, $p < .001$, the RMSEA was .06 [.06, .07] and the CFI was .87. Fit indicators for this model are significantly better than BAPQ-Model indicators although model fit indices cannot be compared directly. Measurement invariance among genders was confirmed also for this model, ΔCFI = .000. On the other hand, the examination of AlRig-Model factor loadings and modification indices is the best way to test if the internal structure problems caused by Pragmatic Language have been now alleviated. Factor loadings and modification indices can be seen in Tables 4 and 5. Factor loadings in absolute values were moderate-high M = .61, range = [.30, .81] for Aloof factor and moderate M = .50, range = [.29, .72] for the Rigid factor. In the same way that factor loadings, Aloof and Rigid omega estimations remained stable (when compared to BAPQ-Model): Omega estimations [95% CI] were .89 [.87, .90] for Aloof dimension and .79 [.77, .82] for Rigid dimension. Correlation between Aloof and Rigid dimensions was − .44 (p < .001). These results support the use of Aloof and Rigid scores for the proposed goals; in this case, only one item (item 19 of Rigid dimension) showed potential cross-loadings.

Table 2

Factor loading, modification indices, and expected parameter change for factor loadings in BAPQ-Model

		Λ	Modification indices	Standardiced expected parameter change	Factor in which the item should load
Aloof	Item 1	.72	10.31	.31	P. Language
	Item 5	-.53			
	Item 9	.76			
	Item 12	.64			
	Item 16	.70	10.00	.31	P. Language
	Item 18	-.51			
	Item 23	-.31			
	Item 25	.69			
	Item 27	-.63			
	Item 28	.51			
	Item 31	-.60			
	Item 36	.80			
Rigid	Item 3	-.50			
	Item 6	.52			
	Item 8	.55	11.00	.21	P. Language
	Item 13	.59			
	Item 15	-.41			
	Item 19	-.58	10.39	.18	Aloof
	Item 22	.72			
	Item 24	.56			
	Item 26	.31			
	Item 30	-.58			
	Item 33	.29			
	Item 35	.35			
Pragmatic language	Item 2	.50	10.24	.39	Aloof
	Item 4	.28			
	Item 7	-.49	45.53	.81	Aloof
	Item 10	<u>.13</u>			
	Item 11	.76			
	Item 14	.43	23.59	.58	Aloof
	Item 17	-.06	15.89	.49	Aloof
	Item 20	<u>.17</u>	14.05	.46	Aloof
	Item 21	<u>-.07</u>	13.22	.25	Rigid
	Item 29	.54			
	Item 32	<u>.03</u>			
	Item 34	-.29			

Note. Non statistically significant factor loadings (λ) were underlined, the rest were statistically significant ($p < .001$)

Table 3

Modification indices pointing potential correlations between items uniqueness in BAPQ-Model

Pair of items	Modification indices	Standardiced expected parameter change	BAPQ factors potentially afected
Item 34-Item 21	62.08	.43	P. Language
Item 32-Item 17	52.19	.39	P. Language
Item 20-Item 14	40.06	.35	P. Language
Item 20-Item 17	36.84	.33	P. Language
Item 17-Item 26	25.36	.27	P. Language and Rigid
Item 14-Item 2	23.73	.28	P. Language
Item 19-Item 16	22.47	.28	Rigid and Aloof
Item 32-Item 14	20.93	.25	P. Language
Item 32-Item 20	20.34	.24	P. Language
Item 17-Item 14	20.27	.25	P. Language
Item 32-Item 26	19.85	.24	P. Language and Rigid
Item 19-Item 6	17.83	-.25	Rigid
Item 9-Item 1	16.23	.26	Aloof
Item 13-Item 23	15.96	.23	Aloof and Rigid
Item 22-Item 19	13.48	.25	Rigid
Item 20-Item 7	12.98	.20	P. Language
Item 34-Item 33	12.82	.20	P. Language and Rigid
Item 29-Item 10	12.55	.20	P. Language
Item 33-Item 19	12.53	.20	Rigid
Item 14-Item 26	11.86	.19	P. Language and Rigid
Item 18-Item 9	11.59	.20	Aloof
Item 14-Item 33	11.56	-.19	P. Language and Rigid
Item 32-Item 35	11.11	.18	P. Language and Rigid
Item 24-Item 23	10.58	.18	Rigid and Aloof
Item 33-Item 24	10.14	.18	Rigid
Item 34-Item 28	10.10	.18	P. Language and Aloof

Table 4

Factor Loading, Modification Indices, and Expected Parameter Change for Factor Loadings in AlRig-Model

		Λ	Modification indices	Standardiced expected parameter change	Factor in which the item should load
Aloof	Item 1	.73			
	Item 5	-.52			
	Item 9	.76			
	Item 12	.62			
	Item 16	.72			
	Item 18	-.51			
	Item 23	-.30			
	Item 25	.67			
	Item 27	-.63			
	Item 28	.51			
	Item 31	-.59			
	Item 36	.81			
Rigid	Item 3	-.50			
	Item 6	.52			
	Item 8	.54			
	Item 13	.59			
	Item 15	-.41			
	Item 19	-.58	11.53	.19	Aloof
	Item 22	.72			
	Item 24	.56			
	Item 26	.31			
	Item 30	-.59			
	Item 33	.29			
	Item 35	.34			

Note. All factor loadings (λ) were statistically significant (p < .001)

Table 5

Modification Indices Pointing Potential Correlations Between Items Uniqueness in AlRig-Model

Pair of items	Modification indices	Standardiced expected parameter change	BAPQ factors potentially afected
Item 19-Item 16	20.74	.27	Rigid and Aloof
Item 19-Item 6	18.14	-.26	Rigid
Item 13-Item 23	16.08	.23	Rigid and Aloof
Item 9-Item 1	14.07	.24	Aloof
Item 22-Item 19	13.44	.25	Rigid
Item 33-Item 19	13.44	.21	Rigid
Item 27-Item 18	10.61	.19	Aloof
Item 18-Item 9	10.54	.20	Aloof
Item 24-Item 23	10.35	.18	Rigid and Aloof

Validity Evidence Based on the Relationships with Measures of Other Variables

Fit indicators for BAPQ-StrModel were: $\chi 2$ (789) = 1925.20, $p < .001$, the RMSEA was .06 [.06, .07] and the CFI was .77. Structural invariance among genders was confirmed for this model, ΔCFI = .003. Since the measurement model included in this model is identical to BAPQ-Model, factor loadings and relationships between uniqueness are the same. Consequently, fit problems are replicated and this model should not even be considered as acceptable. Furthermore, correlations and regression coefficients show certain inconsistent validity evidences when supporting the interpretation of BAPQ factors (see Table 6). As expected, the Pragmatic Language dimension is highly and negatively related with the PAQ subscales, and not related in a statistically significant way with Extraversion, Agreeableness, and Openness. Additionally, Aloof dimension is highly positively related to Extraversion and moderately positively related to Agreeableness, whereas Rigid dimension show low/moderate negative relation with Openness. However, on the other hand, the relationship between Aloof dimension (a high score in this dimension should be interpreted as lack of Aloof) and the PAQ subscales are moderate and negative in this model. These relationships could be taken as slight negative validity evidences.

On the other hand, AlRig-StrModel showed almost acceptable fit estimations $\chi2$ (789) = 854.18, $p < .001$, the RMSEA was .06 [.06, .07] and the CFI was .86 (model fit indices cannot be compared directly with the previous model). As in the measurement model, the structural model shows adequate RMSEA estimations but CFI estimations slightly below the recommended cut-off. Structural invariance among genders was also confirmed for this model, ΔCFI = .003. Correlations, regression coefficients, and R^2 estimations of AlRig-StrModel can be seen in Table 7. This model shows relationships among BAPQ factors and the rest of variables that match with the expected results. Furthermore, in this model the relationships among Aloof and the PAQ subscales also coincide the expected results and the confidence intervals are narrower in contrast to the BAPQ-StrModel.

Table 6

Regression Coefficients [and 95% CI Intervals] for BAPQ-Str Model Above the Diagonal. Correlations and R^2 Below the Diagonal

	Aloof	Rigid	Extraversion	Agreeableness	Openness	Enunciative Pragmatic	Textual Pragmatic	Interactive Pragmatic
Aloof			.82 [.70, .93]**	.39 [.21, .57]**	-.31 [-0.52, -.11]*	-.34 [-.55, -.13]*	-.50 [-.72, -.28]**	-.25 [-.45, -.05]*
Rigid	-.58**		-.12 [-.19, -.04]*	-.01 [-.12, .11]	-.29 [-0.40, -.17]**	-.08 [-.19, .04]	-.10 [-.22, .02]	-.07 [-.17, .04]
Pragmatic Language	-.80**	.33**	.05 [-.09, .18]	.06 [-.15, .27]	-.23 [-0.46, -.01]	-.75 [-.97, -.53]**	-.88 [-1.11, -.65]**	-.86 [-1.06, -.65]**
Extraversion								
Agreeableness			.00					
Openness			.01	.06				
Enunciative Pragmatic			.24**	.02	.15*			
Textual Pragmatic			.08	.06	.26**	.41**		
Interactive Pragmatic			.10	.13	.04	.45**	.51**	
R^2			.70	.12	.10	.32	.38	.50

Note. * $p < .05$; ** $p < .001$.

Table 7

Regression Coefficients [and 95% CI intervals] for AlRig-StrModel Above the Diagonal. Correlations and R^2 Below the Diagonal

	Aloof	**Extraversion**	**Agreeableness**	**Openness**	**Enunciative Pragmatic**	**Textual Pragmatic**	**Interactive Pragmatic**
Aloof		.79 [.74, .83]**	.35 [.25, .44]**	-.15 [-.26,- .04]*	.19 [.08, .29]**	.12 [.02, .22]	.35 [.26, .44]**
Rigid	-.58**	-.11 [-.18,-.04]*	.00 [-.11, .11]	-.33 [-.43,- .22]**	-.21 [-.31, -.1]**	-.25 [-.36, -.14]**	-.21 [-.31, -.11]**
Extraversion							
Agreeableness		.01					
Openness		.00	.06				
Enunciative Pragmatic		.18**	-.00	.20**			
Textual Pragmatic		.04	.03	.30**	.58**		
Interactive Pragmatic		.06	.08	.12*	.60**	.67**	
R^2		.71	.12	.09	.11	.10	.23

Note. * p < .05; ** p < .001.

Discussion

As far as we know, this is the first study that provides a Spanish version of a questionnaire to explore the BAP. Furthermore, this study extends previous findings on the factor structure of the BAPQ (e.g., Broderick et al., 2015; Ingersoll et al., 2011; Sasson, Lam et al., 2013; Wainer et al., 2011) in a large sample of university students. All the obtained results were gender invariant; therefore, gender cannot be considered a variable that may influence them. Our results are in agreement with Hurley et al.'s (2007) internal structure validity evidence for the control group since they found three dimensions and moderate correlations for all the variables (namely, Aloof and Rigid $r = .54$, Aloof and Pragmatic Language $r = .53$, and Rigid and Pragmatic Language $r = .51$), although the correlation between Aloof and Pragmatic Language is rather higher in our study than in the BAPQ original study (Hurley et al., 2007). We consider that a possible severity disparity in the BAP traits, between control parents and university students, could also underlie this BAPQ internal structure difference (e.g., different levels of data dispersion or ceiling/floor effects).

Nevertheless, the current study evidences a bigger problem in the internal structure of the BAPQ. As we hypothesized, the severe misfit of the BAPQ-Model is caused mainly by problems in the Pragmatic Language dimension. The Pragmatic Language items are the main contributors to the global misfit of the BAPQ-Model (BAPQ items 2, 7, 14, 17, 20). In contrast, Aloof and Rigid dimensions show adequate fit with moderate factor loadings and adequate reliability estimations. Our results are similar to those found by Sasson, Lam et al.'s (2013) who indicated that five out of twelve salient cross-loadings were caused by the items of the Pragmatic Language factor (BAPQ items 2, 7, 11, 29, and 34).

The alternative measurement model, excluding the Pragmatic Language dimension, showed CFI estimations slightly below than the recommended cut-off to consider it adequate. However, this CFI estimation is clearly better than that obtained when Pragmatic Language was included in the model as well as than that reported for most of the studies where the BAPQ internal structure was analysed (e.g., Broderick et al., 2015). These results are in agreement with the idea that the main source of misfit in BAPQ dimensions is due to the Pragmatic Language scores. Moreover, we consider that the CFI misfit of the model with only Aloof and Rigid dimensions could be explained by the high ratio item/factor in each BAPQ dimension (see Kenny and McCoach, 2003). These results, along with Aloof and Rigid adequate reliability estimations, support the use of their scores for the proposed goals, although fit indicators suggest that these measures could be improved.

A more detailed analysis of the correlations between item uniqueness in the BAPQ-model (see Table 5) contributes to identify the main potential source of misfit in Aloof and Rigid scales. These results indicate a possible effect of content overlapping, that is, some of the items could be also referred to related contents apart from Aloofness and Rigidity. For example, items 19, 16, and 6 are all referred to cope with new experiences, so these items could be exploring also the preference for this kind of experiences no matter which scale they belong to.

Our second goal was to explore the validity evidences of the BAPQ scores based on its relationship with measures of other variables. In general terms, results from this study indicate that each BAPQ scale is associated to a number of personality traits and language skills as we expected. As it was hypothesized, validity evidences showed a strong negative association between Pragmatic Language and Enunciative Pragmatic, Textual Pragmatic and Interactive Pragmatic from PAQ (Rodríguez-Muñoz, 2012). However, as it was highlighted, we strongly advise against the use of the BAPQ Pragmatic Language factor. It has negative internal structure validity evidences and the inclusion of Pragmatic Language factor in any prediction, together with Aloof, could lead to inaccurate and unexplainable results.

On the other hand, Aloof and Rigid scores have shown empirical evidences supporting their use. In addition to their favourable internal structure validity evidences, Aloof (it should be interpreted as lack of Aloof in the models of the current study) had a strong positive correlation with Extraversion and a moderate-low positive correlation with Agreeableness. Taking into account this information, our second hypothesis should be inverted (because the reverse Aloof interpretation) but accepted. Indeed, the strong correlation between Extraversion and Aloofness could be explained regarding to the negative pole of the Extraversion construct, which also encompasses negative qualifiers like loneliness, self-absorption, social withdrawal, etc., theoretically strongly related to Aloofness. We also found a moderate-low negative association between Rigidity and Openness, which confirm our original hypothesis. According to the Openness definition (see Costa and McCrae, 1999), and Rigidity conceptualization (e.g., Hurley et al., 2007), both constructs should be slight and negatively correlated since they have few qualifiers in common (e.g., Costa and McCrae, 1999).

However, on the other hand, Aloof dimension was moderately and negatively related to the PAQ subscales. These unexpected relationships have to be analysed together with the high Aloof-P. Language correlation and the wide confidence intervals (which sizes are a consequence of standard errors magnitudes) of some of the Aloof and Pragmatic Language regression coefficients. These results could be a direct consequence of collinearity between Aloof and Pragmatic Language factors (O'brien, 2007). This fact is confirmed in the AlRig-StrModel model where an adequate (i.e., positive) relationship between Aloof and PAQ factors with narrower confidence intervals is showed. The results confirmed this hypothesis indicating that the Pragmatic Language malfunctioning not also affect to the internal structure validity evidences, but also to validity evidences of Aloof and Rigid scores based on the relationship with other variables.

Collectively, these results indicate that more research is needed to improve the psychometric properties of this Spanish version of the BAPQ. Nevertheless, Aloof and Rigid subscales can be used with certain guaranties. Our study supports the Sasson, Lam et al.'s (2013) research, which already shows problems in the internal structure of the BAPQ in a sample of parents with ASD children (i.e., they do not achieve a simple factor structure). Consequently, we consider that the psychometric problems we have detected in the questionnaire are not directly derived from the language adaptation process. To sum up, in this study we have presented the first Spanish version of the BAPQ. This test has shown similar psychometric properties than other languages BAPQ versions although our results only support the use of Aloof and Rigid dimensions to assess BAP. Finally, Pragmatic Language scores have shown negative validity evidence similar to those found in other versions of the BAPQ (e.g., Sasson, Lam et al., 2013).

We would like to conclude highlighting that although Aloof and Rigid subscales can be used to study BAP in the Spanish population, it would be reasonable to make efforts to improve BAP measurements regardless of the language version used. As a first step, we propose to run the item response theory

(IRT) approach on BAPQ dimensions; these analyses allow placing items and persons in the same continuous. Thus, IRT would help to identify whether BAPQ items are really suitable to measure BAP in a non-clinical sample (if so, items should be distributed along the same construct levels than this sample). Additionally, we think that the best way to improve BAP measurements could be to construct a new assessment tool based on items which psychometric properties have been widely proven. Based on Wainer et al. (2011) study exploring the common structure underlying AQ, SRS, and BAPQ scores, it would be reasonable to select, among these scales, the items that prove their better psychometric properties for each dimension by using IRT or exploratory factor analysis in the CFA framework (see Brown, 2006). Once a stable subdomain internal structure is achieved, validity evidences could be tested as in the present research but extending the sample to the general population (including not only college students but also people at different ages and with different educational levels) as well as to the relatives of ASD patients.

Acknowledgments. This study was supported by Grant PSI2015- 65248-P from the Spanish Ministry of Economy and Competitiveness, co-funded by ERDF (FEDER) funds. MGG is supported by a predoctoral grant by the Ministry of Education, Culture and Sport (FPU2015-01562). Authors would like to thank to Rocio Rodríguez Robles for her expert English language editing help.

Author Contributions. PSJ and MGG analysed the data and wrote the manuscript. MGG and AGR were responsible for data collection. PSJ, AFE and FC were responsible for the design of the study. All authors participated in the BAPQ adaption into Spanish language, contributed to data interpretation and assisted with writing the manuscript.

Compliance with Ethical Standards.

Conflict of interest. All authors declare no conflict of interest.

References

Amaral, D. G., Dawson, G., & Geschwing, D. H. (2011). *Austism Spectrum Disorders.* New York: Oxford University Press.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®).* American Psychiatric Pub.

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences*, *38*, 451-460. https://doi.org/ 10.1016/j.paid.2004.04.022

Bailey, A., Palferman, S., Heavey, L., & Le Couteur, A. (1998). Autism: The phenotype in relatives. *Journal of Autism and Developmental Disorders, 28,* 369-392. https://doi.org/10.1023/A:1026048320785

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5-17. https://doi.org/ 10.1023/A:1005653411471

Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain research*, *1380*, 42-77. https://doi.org/10.1016/j.brainres.2010.11.078

Bolton, P., Macdonald, H., Pickles, A., Rios, P. A., Goode, S., Crowson, M., ... & Rutter, M. (1994). A case‐control family history study of autism. *Journal of Child Psychology and Psychiatry, 35*, 877-900. https://doi.org/ 10.1111/j.1469-7610.1994.tb02300.x

Broderick, N., Wade, J. L., Meyer, J. P., Hull, M., & Reeve, R. E. (2015). Model Invariance Across Genders of the Broad Autism Phenotype Questionnaire. *Journal of Autism and Developmental Disorders, 45,* 3133-3147. https://doi.org/ 10.1007/s10803-015-2472-z

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York, NY: Guilford Press.

Bruining, H., De Sonneville, L., Swaab, H., De Jonge, M., Kas, M., van Engeland, H., & Vorstman, J. (2010). Dissecting the clinical heterogeneity of autism spectrum disorders through defined genotypes. *PLoS One*, *5*, e10887. https://doi.org/10.1371/journal.pone.0010887

Chen, F. S., & Yoon, J. M. (2011). Brief report: broader autism phenotype predicts spontaneous reciprocity of direct gaze. *Journal of Autism and Developmental Disorders, 41*, 1131-1134. https://doi.org/ 10.1007/s10803-010-1136-2

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling*, *9*, 233–255. https://doi.org/10.1207/s15328007sem0902_5

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Archives of General Psychiatry, 60,* 524-530. https://doi.org/ 10.1001/archpsyc.60.5.524

Constantino, J. N., & Todd, R. D. (2005). Intergenerational transmission of subthreshold autistic traits in the general population. *Biological Psychiatry,* 57, 655-660. https://doi.org/10.1016/j.biopsych.2004.12.014

Constantino, J. N., & Gruber, C. P. (2007). *Social responsiveness scale (SRS).* Los Angeles, CA: Western Psychological Services.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4,* 5. https://doi.org/10.1037/1040-3590.4.1.5

Costa, P. T., & McCrae, R. R. (1999). *Inventario de Personalidad NEO revisado (NEO PI-R) e Inventario NEO reducido de cinco factores (NEO-FFI) [NEO PI-R, Revised NEO Personality Inventory and NEO Five-Factor Inventory (NEO-FFI)].* Madrid: TEA Editions.

Dawson, G., Estes, A., Munson, J., Schellenberg, G., Bernier, R., & Abbott, R. (2007). Quantitative assessment of autism symptom-related traits in probands and parents: Broader phenotype autism symptom scale. *Journal of Autism and Developmental Disorders, 37*(3), 523–536. https://doi.org/10.1007/s10803-006-0182-2

D'Cruz, A. M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., & Sweeney, J. A. (2013). Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology, 27,* 152. https://doi.org/10.1037/a0031721

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. https://doi.org/ 10.3758/brm.41.4.1149

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research, 65,* 591-598. http://doi.org/ 10.1203/PDR.0b013e31819e7203

Ganz, J. B. (2007). Classroom structuring methods and strategies for children and youth with autism spectrum disorders. *Exceptionality, 15,* 249-260. https://doi.org/10.1080/09362830701655816

Geschwind, D. H. (2009). Advances in autism. *Annual Review of Medicine, 60,* 367-380. https://doi.org/10.1146/annurev.med.60.053107.121225

Grinter, E. J., Van Beek, P. L., Maybery, M. T., & Badcock, D. R. (2009). Brief report: Visuospatial analysis and self-rated autistic-like traits. *Journal of Autism and Developmental Disorders, 39,* 670-677. https://doi.org/ 10.1007/s10803-008-0658-3

Hair, J. F., Black, W. C., Babin, B. J., & Anderson R. E. (2009). *Multivariate Data Analysis: A Global Perspective (7th ed.).* Upper Saddle River: Prentice Hall.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. https://doi.org/10.1080/10705519909540118

Hurley, R. S., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders, 37,* 1679-1690. https://doi.org/ 10.1007/s10803-006-0299-3

Ingersoll, B. (2010). Broader autism phenotype and nonverbal sensitivity: evidence for an association in the general population. *Journal of Autism and Developmental Disorders, 40,* 590-598. https://doi.org/ 10.1007/s10803-009-0907-0

Ingersoll, B., Hopwood, C. J., Wainer, A., & Donnellan, M. B. (2011). A comparison of three self-report measures of the broader autism phenotype in a non-clinical sample. *Journal of Autism and Developmental Disorders, 41,* 1646-1657. https://doi.org/10.1007/s10803-011-1192-2

Jeste, S. S., & Geschwind, D. H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nature Reviews Neurology, 10,* 74-81. https://doi.org/ 10.1038/nrneurol.2013.278

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2,* 217-250.

Kenny, D. A., & McCoach, D. B. (2003). Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 10*(3), 333–351. https://doi.org/10.1207/s15328007sem1003_1

Kenny, D. A., & McCoach, D. B. (2003). Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 10*, 333–351. https://doi.org/10.1207/s15328007sem1003_1

Klusek, J., Losh, M., & Martin, G. E. (2014). Sex differences and within-family associations in the broad autism phenotype. *Autism, 18*, 106-116. https://doi.org/ 10.1177/1362361312464529

Landry, O., & Chouinard, P. A. (2016). Why we should study the broader autism phenotype in typically developing populations. *Journal of Cognition and Development, 17*, 584-595. https://doi.org/10.1080/15248372.2016.1200046

Lenroot, R. K., & Yeung, P. K. (2013). Heterogeneity within autism spectrum disorders: what have we learned from neuroimaging studies? *Frontiers in Human Neuroscience, 7*. https://doi.org/10.3389/fnhum.2013.00733

Losh, M., Childress, D., Lam, K., & Piven, J. (2008). Defining key features of the broad autism phenotype: A comparison across parents of multiple-and single-incidence autism families. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 147,* 424-433. https://doi.org/10.1002/ajmg.b.30612

Losh, M., Adolphs, R., Poe, M. D., Couture, S., Penn, D., Baranek, G. T., & Piven, J. (2009). Neuropsychological profile of autism and the broad autism phenotype. *Archives of General Psychiatry, 66,* 518-526. https://doi.org/ 10.1001/archgenpsychiatry.2009.34

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84–99. https://doi.org/ 10.1037/1082-989x.4.1.84

McCrae, R. R., & Costa, P. T. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of Personality and Social Psychology, 56*, 586. https://doi.org/10.1037/0022-3514.56.4.586

McDonald, R. P. (1999). *Test theory: A unified treatment.* Mahwah, NJ: Erlbaum.

Meera, S. S., Girimaji, S. C., Seshadri, S. P., Philip, M., Shivashankar, N., Morgan, P., & Piven, J. (2015). Translation of the Broad Autism Phenotype Questionnaire to an Indian language: A description of the process. *Asian Journal of Psychiatry, 15*, 62–67. https://doi.org/ 10.1016/j.ajp.2015.04.013

Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Guidelines for test translation and adaptation: second edition [Directrices para la traducción y adaptación de los tests: segunda edición]. *Psicothema, 25*, 151-157.

Muñoz, F. J. R. (2012). La conciencia pragmática de adultos con síndrome de Asperger. *Revista de Logopedia, Foniatría y Audiología, 32*, 21-31. https://doi.org/10.1016/j.rlfa.2012.01.002

O'brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity, 41*, 673–690. https://doi.org/ 10.1007/s11135-006-9018-6

Pisula, E., & Ziegart-Sadowska, K. (2015). Social Communication and Language Deficits in Parents and Siblings of Children with ASD–a Short Review. *Autism.* Rijeka: InTech. https://doi.org/ 10.5772/59134

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997). Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry, 154*, 185-190. https://doi.org/10.1176/ajp.154.2.185

Piven, J., Palmer, P., Landa, R., Santagelo, S., Jacobi, D., & Childress, D. (1997). Personality and language characteristics in parents from multiple-incidence autism families. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 74*, 398-411. https://doi.org/10.1002/(sici)1096-8628(19970725)74:4<398::aid-ajmg11>3.0.co;2-d

Piven, J., Wzorek, M., Landa, R., Lainhart, J., Bolton, P., Chase, G. A., & Folstein, S. (1994). Personality characteristics of the parents of autistic individuals. *Psychological medicine*, *24*, 783-795. https://doi.org/10.1017/S0033291700027938

Prutting, C. A., & Kittchner, D. M. (1987). A clinical appraisal of the pragmatic aspects of language. *Journal of Speech and Hearing Disorders, 52*, 105-119. https://doi.org/ 10.1044/jshd.5202.105

Raykov, T., & Marcoulides, G. A. (2004). Using the Delta Method for Approximate Interval Estimation of Parameter Functions in SEM. *Structural Equation Modeling: A Multidisciplinary Journal, 11*, 621–637. https://doi.org/10.1207/s15328007sem1104_7

Raykov, T., & Marcoulides, G. A. (2011). *Introduction to Psychometric Theory.* New York, NY: Routledge.

Ronald, A., Happe, F., Bolton, P., Butcher, L. M., Price, T. S., Wheelwright, S., ... & Plomin, R. (2006). Genetic heterogeneity between the three components of the autism spectrum: a twin study. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*, 691-699. https://doi.org/10.1097/01.chi.0000215325.13058.9d

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism, 6,* 2. https://doi.org/10.1186/2040-2392-6-2

Sasson, N. J., Lam, K. S. L., Childress, D., Parlier, M., Daniels, J. L., & Piven, J. (2013). The Broad Autism Phenotype Questionnaire: Prevalence and Diagnostic Classification. *Autism Research, 6*, 134–143. https://doi.org/ 10.1002/aur.1272

Sasson, N. J., Nowlin, R. B., & Pinkham, A. E. (2013). Social cognition, social skill, and the broad autism phenotype. *Autism, 17*, 655-667. https://doi.org/ 10.1177/1362361312455704

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *MPR-online, 8*, 23-74.

Shi, L.-J., Ou, J.-J., Gong, J.-B., Wang, S.-H., Zhou, Y.-Y., Zhu, F.-R., ... Luo, X.-R. (2015). Broad autism phenotype features of Chinese parents with autistic children and their associations with severity of social impairment in probands. *BMC Psychiatry, 15*, 168. https://doi.org/ 10.1186/s12888-015-0568-9

Stewart, M. E., & Austin, E. J. (2009). The structure of the Autism-Spectrum Quotient (AQ): Evidence from a student sample in Scotland. *Personality and Individual Differences, 47*, 224-228. https://doi.org/10.1016/j.paid.2009.03.00 4

Stewart, M. E., & Ota, M. (2008). Lexical effects on speech perception in individuals with "autistic" traits. *Cognition, 109*, 157-162. https://doi.org/10.1016/j.cognition.2008.07.010

Sucksmith, E., Roth, I., & Hoekstra, R. A. (2011). Autistic traits below the clinical threshold: re-examining the broader autism phenotype in the 21st century. *Neuropsychology review, 21*, 360-389. https://doi.org/10.1007/s11065-011-9183-9

Thurstone, L. L. (1947). Factorial analysis of body measurements. *American Journal of Physical Anthropology, 5*, 15-28. https://doi.org/10.1002/ajpa.1330050103

Wainer, A. L., Ingersoll, B. R., & Hopwood, C. J. (2011). The Structure and Nature of the Broader Autism Phenotype in a Non-clinical Sample. *Journal of Psychopathology and Behavioral Assessment, 33*, 459–469. https://doi.org/10.1007/s10862-011-9259-0

Wakabayashi, A., Baron-Cohen, S., & Wheelwright, S. (2006). Are autistic traits an independent personality dimension? A study of the Autism-Spectrum Quotient (AQ) and the NEO-PI-R. *Personality and Individual Differences, 41*, 873-883. https://doi.org/10.1016/j.paid.2006.04.003

CHAPTER FOUR. STUDY 2:

APPLYING A RASCH MODEL TO THE BAPQ: ITEM SEVERITY ANALYSIS AND TEST DIFFERENTIAL FUNCTIONING OF THE ENGLISH AND SPANISH VERSIONS

Manuscript submitted as:

Godoy-Giménez, M., González-Rodríguez, A., Estévez, A. F., Cañadas, F., & Sayans-Jiménez, P. (2018). Applying a Rasch Model to the BAPQ: Item Severity Analysis and Test Differential Functioning of the English and Spanish Versions. *Current Psychology*.

Applying a Rasch Model to the Broad Autism Phenotype Questionnaire: Item Severity Analysis and Test Differential Functioning of the English and Spanish Versions

Marta Godoy-Giménez[1], Antonio González-Rodríguez[1], Angeles F. Estévez[1, 2], Fernando Cañadas[1, 2], and Pablo Sayans-Jiménez[1]

[1]Department of Psychology, University of Almería

[2]Cernep Research Centre, University of Almería

Author Note

Marta Godoy-Giménez https://orcid.org/0000-0002-2298-0113

Antonio González-Rodríguez https://orcid.org/0000-0001-6432-829X

Angeles F. Estévez http://orcid.org/0000-0003-0067-0840

Fernando Cañadas http://orcid.org/0000-0001-7689-0132

Pablo Sayans-Jiménez https://orcid.org/0000-0001-6681-5729

Correspondence concerning this article should be addressed to Pablo Sayans-Jiménez, Department of Psychology, University of Almería, 04120 Almería, Spain. Email: psj242@ual.es and Fernando Cañadas, Department of Psychology, Cernep Research Centre, University of Almería, 04120 Almería, Spain. Email: jcanadas@ual.es. Funding. This work was funded by Grant PID2019-110066GB-I00 from the Spanish Ministry of Science and Innovation, co-funded by ERDF (FEDER) funds; and Grant FPU2015-015 from the Spanish Ministry of Education, Culture, and Sport. Authors have no known conflict of interest to disclose.

Abstract

The broad autism phenotype (BAP) represents the elevated but nonclinical levels of autism-related expressions expanding them beyond the threshold disorder towards the general population and it is assessed worldwide using the Broad Autism Phenotype Questionnaire. In this study, we applied a Rasch item response approach to the Spanish and the English BAPQ versions. This approach allows us to test the proper functioning of each version, estimate which BAP behaviours are more likable at different levels of severity, and study whether BAPQ item's severities and orders could be considered equivalent between both versions. A Spanish community sample of 970 participants and an English of 533 ones completed either the Spanish or the English version of the BAPQ. The results revealed a lack of unidimensionality of the Pragmatic Language subscale in both test versions and the Rigid subscale in the English version. Both Aloof subscale versions and the Spanish Rigid demonstrated adequate properties but with several items showing differential functioning in the case of Aloof. We conclude by highlighting the necessity of BAP measurements paralleling the current autism spectrum disorder structure, following the severity-dimensional conceptualization (behaviours are continuously distributed according to their severity), and suitable for being adapted to different languages.

Keywords: Broad Autism Phenotype; BAPQ; Item Response Theory; Rating Scale Model; Autism Developmental Disorders

Applying a Rasch Model to the Broad Autism Phenotype Questionnaire: Item Severity Analysis and Test Differential Functioning of the English and Spanish Versions

The autism spectrum hypothesis proposes that autistic behaviours and personality expressions would be presented not only in clinical populations but rather within and along with the general population in different severity (Hoekstra et al., 2008). Inside this spectrum, the broad autism phenotype (BAP) is defined by elevated but nonclinical levels of autism-related symptoms expanding them beyond the threshold of the disorder towards the general population levels (De Groot & Van Strien, 2017). Thus, the BAP expressions inserted in the autism spectrum would contribute to the understanding of the disorder and its developmental trajectories (Landry & Chouinard, 2016). Also, the BAP opens the possibilities of carrying out studies with larger and more heterogeneous samples and, so, entailing a better representation of the population and guaranteeing greater statistical power (Mitchell & Jolley, 2013).

When studying the BAP, the Broad Autism Phenotype Questionnaire (BAPQ; Hurley et al., 2007) has been extendedly applied over different countries and cultures (e.g., *China*, Shi et al., 2015; *Israel,* Seidman et al., 2012). It was designed to measure the BAP in parents of children diagnosed with autism spectrum disorder (ASD; APA, 2013) and it has been subsequently applied to broader community samples (e.g., Álvarez-Couto et al., 2021; Godoy-Giménez et al., 2018; Jamil et al., 2017; Jakobson et al., 2018; Morrison et al., 2018; Stojković et al., 2018). It represents the BAP original structure: Aloof personality, Rigid personality, and Pragmatic Language impairment (Piven et al., 1997a; Piven et al., 1997b) which paralleled autism disorder domains of impairment of the former DSM-IV ASD operationalization (APA, 1994).

While research in psychometric properties of the BAPQ has provided validity evidence on the BAPQ scores' inferences (based on their relations with other variables, Sasson et al., 2013b; on test-criterion relationships, Broderick et al., 2015; on differential scores among groups, Shi et al., 2015; and on comparing BAPQ versions, Sasson et al., 2014), little is known about BAPQ reliability at the different levels of the severity continuum, item severities or the adequacy of response categories. According to the autism spectrum hypothesis, understanding that autism expressions are spread somehow across the population implies that the items of the BAPQ could be scaled according to their severity. It would imply that the items conforming to each of the three BAPQ subscales should express different degrees of severity to grasp the essence of the spectrum of autism (at subclinical levels in the case of the BAP) and, at the same time, that every person may be located at some point inside this continuum even those who express mild severity BAP behaviours.

Item response theory (IRT) become of great relevance in this respect as they could bring many advantages in the study of BAPQ psychometric properties (e.g., Bond & Fox, 2015) providing, among others: (i) both interval-level scaling of persons and precision estimations for each severity level (essential for interpreting BAPQ's scores according to their severity), (ii) test-independent scores, useful for comparing BAP levels measured with different tests, or (iii) conjoint scaling of items and respondents along the continuum and for inferring which behaviours are more likable at different levels of severity. This conjoint estimation will also allow inferring the amount of BAP severity of each item (the magnitude of BAP that would be required to adhere to each expression, scaled in a severity continuum from above-average functional behaviours to severe autistic-associated behaviours which require substantial help) and will ease the discussion about the differential performance of the items in different groups (e.g., ASD relatives vs. the general population). Furthermore, the resulting severity order can be used as an additional source of validity evidence.

On the other hand, many authors have highlighted the importance of assessing differential test functioning (DTF) and differential item functioning (DIF) to assure that all the test adaptions preserve the equivalence of measurement (item hierarchies and/or severities) among cultures (Hambleton, 1994). Unlike other questionnaires that have analysed test stability between versions (the Social Responsiveness Scale—SRS; Bölte et al., 2008), the BAPQ DFT has never been tested across test adaptations. This is especially important considering that some social expressions linked to ASD vary their severity depending on the country (Kim, 2012).

This study aims to further analyse the psychometric properties of the BAPQ through IRT models. The importance of this study relies on the ASD dimensional approach in the DSM-V (APA, 2013) and the autism spectrum hypothesis (De Groot

& Van Strien, 2017) summed to the necessity of exploring whether the BAP operationalization targeted in the BAPQ would follow this severity-dimensional approach. That is, whether the items of the three subscales of the BAPQ could be graded in severity and displayed creating three hierarchies of severity in three different domains (corresponding to Aloof, Pragmatic Language, and Rigid subscales).

For those purposes, the original BAPQ (Hurley et al., 2007) and its Spanish version, the BAPQ-SP (Godoy-Giménez et al., 2018) will be applied to two community samples (one from the United Kingdom and one from Spain) and analysed using a Rasch rating scale model (RSM, Masters & Wright, 1984).

Secondly, we will test whether the items' hierarchies and severity of each BAPQ item remain stable in its Spanish version. DTF (equivalence in the functioning of sets of items) and DIF (the loss of item estimate invariance across subsamples of respondents; Bond & Fox, 2015) between both BAPQ versions will reveal that the members of different groups with the same level of BAP would answer in significantly different ways to the same items (e.g., Hambleton, 2006). Even if the severities of some autism-related characteristics could be affected by cultural idiosyncrasies, item invariance failures could also alert us to potential problems with the test (Bond & Fox, 2015). If so, we will consider whether these severity differences are due to cultural differences. This aim is also very important since the BAPQ is one of the most translated tests of the BAP (e.g., Bang et al., 2021 [Swedish version]; Godoy-Giménez et al., 2018 [Spanish version]; Shi et al., 2015 [Chinese version]) and no evidence has been documented about the invariance structure of the BAPQ either the functioning of its items across versions.

The following results would support the use of the original BAPQ and the BAPQ-SP: (i) an adequate item fit informing about the capability of the resulting item hierarchy to predict people patterns of responses (i.e., items statements should endorse progressively from the least to the most severe); (ii) separation indices indicating that it is possible to distinguish, at least, between two statistically different strata in the sample; (iii) item-person maps showing the BAPQ items are distributed covering the severity levels of the persons, and (iv) all the response categories being distinguishable for the respondents. We believe that our results will be in line with previous studies indicating adverse evidence based on the BAPQ internal structure (e.g., Lin et al., 2021; Stojković et al., 2018). Especially, a lack of unidimensionality of some BAPQ subscales is expected, mainly in the case of the Pragmatic Language or Rigid subscales (see Godoy-Giménez et al., 2018; Lin et al., 2021; Sasson et al., 2013a; Sharma et al., 2018; Stojković et al., 2018). If results matched with expectations, it would be impossible to perform RSM in these subscales. Additionally, we expect to find some item severity equivalence among the BAPQ items when comparing the English (BAPQ-EN) and the Spanish (BAPQ-SP) versions. Finally, but not less important, in this study we also present the correction of two items of the Spanish BAPQ (items 4 and 23; Godoy-Giménez et al., 2018). Precedent Spanish version of item 4 does not completely represent the original content of the item in the BAPQ-EN (the Spanish item 4 is expressing the idea of *following the track of the general conversation* while the original item 4 in the BAPQ-EN is expressing *keeping the track of your point in a conversation*). For item 23, it Spanish version seems to express a preference for superficial conversations (an aloof prototypical behavior; corresponding to a direct item) whereas in the BAPQ-EN this item reflects the idea of being good at casual or spontaneous conversations (reversed item). Alternative translations for these items are provided (see instruments section). After adjusting both translations, we will expect that they will not exhibit adverse item fit. A closer inspection of those items will be performed as we expect that alternative translations will work adequately and they will not generate any discrepancy in their severity estimation and comparisons between both BAPQ versions.

Method

Participants

Quota sampling was performed for having access to representative samples, regarding age and sex, of the Spanish and English populations. In the case of Spain, seven quotas by age and gender were calculated on a sample of 600 participants. Through incidental snowball, university students helped us to have access to a larger and general sample that involved themselves, relatives, friends, and acquaintances (undergraduate received two-course credits for collaborating). Later on, incidental sampling was conducted until reaching 970 Spanish participants (Social Networks and Spanish associations of autism took part). The English sample was recruited from an international company (https://www.gfk.com) and it consisted of 533 participants. We prepared eight quotas on a sample of 500 participants and we gave them to the company. The enterprise posted a notice with the study information in an online panel of respondents and participants who met the criteria were given access to the survey. Accessibility was restricted once a quota was completed. Participants received €6 for their collaboration. We took nationality as an exclusion criterion in both cases.

Sample sizes adequacy was calculated for item calibrations. Considering a test with 36 polytomous items, a sample of 500 would produce robust statistically stable measures (item calibrations or person measures stables within ±1.0 logits, robust confidence) in adverse circumstances (Aziza et al.., 2020; Linacre, 2002; Wright & Stone, 1979). Furthermore, in the present study item separation indices were tested to confirm that both sample sizes were adequate to estimate the item hierarchies (as exposed hereafter, indices > 3 assume that the sample size is large enough to verify the item severity hierarchy; Linacre, 2014; and to find the same item placements if identical items are applied to another same-sized sample behaving similarly; Bond & Fox, 2015). Finally, for conducting comparisons between groups (testing differential item or test functioning) and for longer than two-item scale, a sample size of 200 participants per group is required for adequate power (>80%; Scott et al., 2009).

Instruments

The BAPQ self-report form (Hurley et al., 2007) is a 36-item screening questionnaire that demands participants to answer how frequently a statement applies to them on a 6-point scale from 1 "*Very rarely*" to 6 "*Very often*". BAPQ items are grouped into three subscales: aloof personality, rigid personality, and pragmatic language problems (items conforming to each subscale and reversed items can be found in Hurley et al., 2007). The self-report BAPQ-SP (Godoy-Giménez et al., 2018) was applied in this study together with the original English self-report one. Alternatives translations of items 4 (*"Me cuesta evitar irme por las ramas en una conversación"*) and 23[i] (*"Se me da bien la charla insustancial (estar de cháchara)"*) was included this time.

Procedure

The data was compiled in a single session of approximately 30 minutes; both samples had equal conditions. A survey was created in the survey online administrator LimeSurvey (https://www.limesurvey.org/) and it includes the BAPQ together with another test out of scope of the present study and some sociodemographic questions always displayed at the end. The survey had two versions, the English-based language version and the Spanish one. A link drove each participant to an online and completely anonymous survey (links were different depending on the survey language). Personal links (anonymous tokens) were created for the English sample as the company required it to provide a refund for participants' collaboration. All the participants read the study instructions, the data treatment ethics, and expressed their consent to participate in the study and to the use of their data only for scientific purposes before starting. They were asked to read each question carefully and answer them sincerely. They were also warned that aleatory patterns of responses and short time lapses will be tracked and were reasons enough to remove collaboration reinforcers (further information about the procedure can be consulted in MASKED authors., 2018).

Data analysis

Descriptive statistics and independent sample t-tests for continuous variables were calculated to compare both BAPQ versions' means (statistical significance was set at p < .001 using Bonferroni correction). Effect sizes were estimated with Cohen's d using G*Power software v3.1 (d = 0.2 small effect size, 0.5 medium effect size, and 0.8 large effect size; Cohen, 1988). Dimensionality was checked using a principal components analysis of residuals (its interpretation differs from dimensionality studies using factor analysis). If unidimensionality was achieved, Rasch rating scale models (RSM[8]) for tests that have the same rating scales in all items were applied to each subscale separately (6-option Likert in the BAPQ; see Wright & Masters, 1982). RSM analyses were conducted in WINSTEPS software version 3.63.2 (Linacre, 2018) and plots were drawn with R software version 3.6.1 (R Core Team, 2019).

[8] RSM assumes that the distance among every response category will be the same for all categories and items. It allows explore the essential properties of the BAPQ's subscales in six stages (corresponding with results subsections).

In the first place, item separation indices were tested to confirm that the sample size was adequate to estimate the item hierarchy. Indices over 3 assume that the person sample is large enough to verify the item severity hierarchy (Linacre, 2014) and to consider that it is highly likeable to find the same item placements if identical items are applied to another same-sized sample behaving similarly (Bond & Fox, 2015). In a second place, item fit (i.e., the degree to which the proposed item location can predict consistently participants' responses to these items) was analyzed (see Smith et al., 2008). Both, infit (sensitive to patterns of response of people with similar severity levels than one of the items) and outfit (sensitive to patterns of response of people with severity levels far from one of the items) mean-squared residual differences were estimated (values between 0.6 and 1.5 are interpreted as acceptable item fit; Lunz et al., 1990; Wright & Masters, 1982). In a third place, the locations of the items along each severity continuum were analyzed according to their content and the coherence of the item distribution was taken as a source of construct validity. In fourth place, the category probability curves of the items were analyzed. We explored whether the categories were ordered as expected and, whether they had the highest probability of adhesion, at least, at one point of the continuum. In fifth place, the efficacy of BAPQ to distinguish people along the different severity continuums was studied through people's separation indices and person reliability (i.e., the proportion of observed variance not due to measurement error; equivalent to the traditional test reliability). Person separation indices lower than 2 (person reliability < .80) indicate that the test cannot even distinguish between high and low severity levels (Bond & Fox, 2015). Furthermore, the precision of each set of items (i.e., Aloof, Rigid, and Pragmatic Language) throughout their corresponding severity continuum was examined using their test information functions (subscale in this case). Finally, we compared the performance between both versions of each subscale as depicted by their items; that is, DFT between the subscales of the English and the Spanish BAPQs with identity plots (for the whole analysis, consult Bond & Fox, 2015; see Figure 5). These scatterplots represent item severity of each subscale version in each axis, an identity line symbolizing the expected values for severity invariance, and a pair of 95% quality control lines.

Results

The BAPQ-EN items showed higher averages and standard deviations than items in the BAPQ-SP (see Table 1 and Figure 1). Particularly, nine, 10, and 11 items showed higher means in the Aloof, Pragmatic Language, and Rigid subscales respectively. The means of the items in the BAPQ-EN as well as in the BAPQ-SP were low, some of them were between 3-4 and only two items in the BAPQ-SP were beyond 4.

Twenty-four items showed significant differences between both samples (significance was set at $p < .001$ using Bonferroni correction). Cohen's *d* revealed 22 items with small effect sizes (effect sizes; d = 0.2; items 1[i], 2, 3[i], 5, 6, 8, 9[i], 10, 11, 13, 14, 18, 19[i], 21[i], 24, 25[i], 28[i], 29, 30[i], 31, 33, 35), two mediums (d = 0.5; items 4 and 23[i]) and none item with large effect sizes (d = 0.8).

Table 1

Item Descriptive Statistics and IRT Parameters

Item	**M (SD)**		**Sk**		**K**		**Measure**		**SE**		**Infit**		**Outfit**		**Item-total correlation**		**t-test**		
Aloof	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	*t*	*p*	*d*
1[i]	3.30 (1.36)	2.74 (1.30)	0.03	0.39	-0.62	-0.60	-0.07	0.26	.04	.03	0.64	0.78	0.63	0.80	.81	.75	7.71	.00	.42
5	3.25 (1.29)	2.95 (1.39)	0.30	0.52	-0.53	-0.48	-0.02	0.02	.04	.03	1.20	1.04	1.26	1.11	.57	.70	-4.23	.00	.22
9[i]	3.40 (1.46)	3.09 (1.46)	0.06	0.26	-0.81	-0.81	-0.17	-0.13	.04	.03	0.69	0.80	0.68	0.81	.82	.79	-0.97	.00	.21
12[i]	2.91 (1.34)	3.04 (1.23)	0.31	0.26	-0.71	-0.91	0.34	-0.08	.05	.03	1.08	1.22	1.11	1.28	.65	.67	1.74	.08	.10
16[i]	3.14 (1.25)	3.67 (1.47)	0.03	-0.15	-0.60	-0.91	-0.47	-0.75	.04	.03	0.80	0.87	0.81	0.88	.76	.77	-0.46	.65	.39
18	3.25 (1.23)	2.83 (1.29)	0.24	0.65	-0.12	0.02	-0.01	0.16	.04	.03	1.45	1.25	1.65	1.41	.43	.59	-6.15	.00	.33
23[i]	3.54 (1.46)	2.83 (1.27)	0.01	0.66	-0.92	0.03	-0.30	0.16	.04	.03	1.17	1.32	1.18	1.36	.66	.57	-9.39	.00	.52
25[i]	3.54 (1.46)	3.11 (1.38)	0.01	0.29	-0.92	-0.65	-0.31	-0.16	.04	.03	0.84	0.86	0.83	0.86	.73	.75	-5.82	.00	.30
27	2.71 (1.32)	2.62 (1.42)	0.59	0.87	-0.07	-0.02	0.56	0.41	.05	.04	1.25	0.93	1.21	0.87	.60	.75	-1.27	.21	.07
28[i]	2.65 (1.23)	2.85 (1.39)	0.41	0.48	-0.55	-0.55	0.63	0.14	.05	.03	1.09	1.04	1.02	1.11	.62	.70	2.86	.00	.15
31	3.56 (1.41)	3.01 (1.32)	0.24	0.63	-0.68	-0.01	-0.33	-0.05	.04	.03	1.13	0.98	1.15	1.04	.65	.68	-7.35	.00	.40
36[i]	3.09 (1.42)	2.96 (1.25)	0.14	0.31	-0.81	-0.70	0.15	0.01	.04	.03	0.68	0.64	0.66	0.66	.82	.82	-1.74	.08	.10
PL	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	*t*	*p*	*d*

Item	M (SD)		Sk		K		Measure		SE		Infit		Outfit		Item-total correlation		t-test		
2	2.88 (1.37)	2.52 (1.34)	0.48	0.89	-0.35	0.14											-4.92	.00	.27
4	3.00 (1.21)	2.12 (1.08)	0.48	1.18	0.07	1.42											-13.94	.00	.69
7[i]	3.13 (1.25)	3.09 (1.25)	0.12	0.18	-0.47	-0.54											-0.64	.52	.03
10	2.25 (1.35)	1.89 (1.09)	0.96	1.43	0.18	2.02											-5.55	.00	.29
11	2.69 (1.43)	2.50 (1.37)	0.54	0.91	-0.55	0.17											-2.48	.00	.14
14	2.74 (1.29)	2.33 (1.21)	0.57	0.95	-0.12	0.58											-6.13	.00	.33
17	2.48 (1.33)	2.55 (1.41)	0.71	0.81	-0.19	-0.17											0.88	.37	.05
20	3.06 (1.33)	2.85 (1.53)	0.40	0.54	-0.38	-0.75											-2.85	.01	.15
21[i]	3.03 (1.23)	2.76 (1.24)	0.12	0.40	-0.54	-0.47											-4.06	.00	.22
29	2.88 (1.21)	2.35 (1.05)	0.57	0.99	0.05	1.32											-8.56	.00	.47
32	2.89 (1.25)	2.86 (1.40)	0.51	0.59	-0.14	-0.37											-0.38	.70	.02
34[i]	3.14 (1.18)	3.01 (1.25)	0.09	0.31	-0.47	-0.54											-2.06	.04	.11
Rigid	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	EN	SP	*t*	*p*	*d*
3[i]	3.54 (1.31)	4.01 (1.29)	-0.17	-0.38	-0.60	-0.53		-1.01		.03		1.11		1.15	_	.58	-6.64	.00	.36
6	3.05 (1.28)	2.54 (1.29)	0.41	0.86	-0.27	0.30		0.46		.03		1.05		1.04	_	.64	-7.35	.00	.40

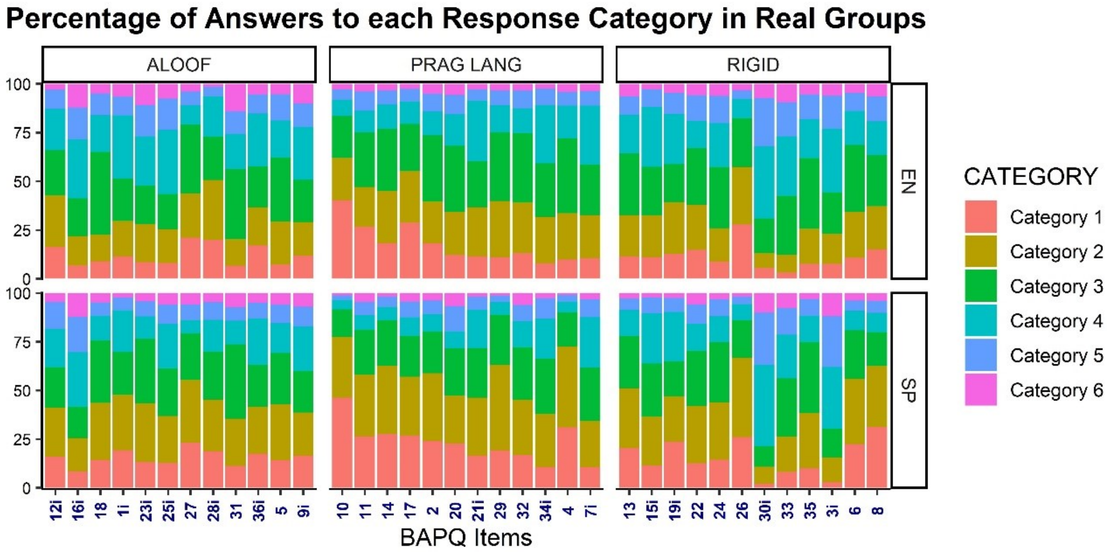
Item	M (SD)		Sk		K		Measure	SE	Infit	Outfit	Item-total correlation	t-test		
8	3.10 (1.44)	2.41 (1.39)	0.30	0.94	-0.77	0.10	0.61	.03	1.23	1.15	_ .65	-8.99	.00	.49
13	3.14 (1.34)	2.62 (1.27)	0.34	0.65	-0.45	-0.05	0.36	.03	0.85	0.83	_ .69	-7.39	.00	.40
15[i]	3.14 (1.25)	3.01 (1.23)	0.03	0.17	-0.60	-0.56	-0.04	.03	0.91	0.91	_ .63	-1.90	.06	.10
19[i]	3.09 (1.37)	2.77 (1.38)	0.23	0.28	-0.79	-0.89	0.21	.03	1.17	1.15	_ .65	-4.37	.00	.23
22	3.05 (1.42)	2.97 (1.37)	0.37	0.55	-0.70	-0.40	0.00	.03	0.83	0.82	.73	-1.11	.27	.06
24	3.34 (1.32)	2.82 (1.27)	0.13	0.55	-0.56	-0.20	0.15	.03	0.74	0.77	.72	-7.45	.00	.40
26	2.44 (1.29)	2.29 (1.16)	0.86	1.05	0.30	0.93	0.75	.04	1.25	1.26	.50	-2.20	.03	.12
30[i]	3.90 (1.23)	4.13 (1.14)	-0.53	-0.55	-0.00	0.21	-1.13	.03	0.83	0.85	.62	3.52	.00	.19
33	3.79 (1.21)	3.39 (1.35)	-0.00	0.16	-0.35	-0.63	-0.41	.03	1.21	1.24	.56	-5.85	.00	.31
35	3.29 (1.27)	2.92 (1.20)	0.25	0.54	-0.39	-0.01	0.05	.03	0.91	0.96	.60	-5.47	.00	.30

Note. Sk: skewness; K: kurtosis; Measure: location of each item along the continuum; SE: standard error; Infit/Outfit: mean squared residual. The significance level was fixed at <.001 using Bonferroni correction.

[i] Inverted items.

Figure 1

Item Descriptive Statistics: Percentage of Responses to Each Response Category in Real Groups



1. Subscale dimensionality and item separation indices

Results pointed to unidimensionality of both Aloof subscales (variances explained by measures were BAPQ-EN: 50.60% and BAPQ-SP: 55.00%). The BAPQ-EN Rigid subscale seemed to present lack of unidimensionality (variance explained by measure: 45.3%, eigenvalue = 2.769) with a secondary dimension (explained variance in the first contrast 12.60%, eigenvalue = 2.76, disattenuated correlation with de primary dimension = .41) formed by items 3^i, 15^i, 19^i, and 30^i. The BAPQ-SP Rigid subscale was unidimensional (variance explained by measures: 51.10%). Finally, lack of unidimensionality was also observed in both Pragmatic Language subscales (explained variances by measures were BAQ-EN: 37.90%, BAPQ-SP: 42.00%). In the case of the BAPQ-EN, the secondary dimension (explained variance in first contrast 14.2%, eigenvalues = 2.74, disattenuated correlation with the primary dimension = .21) was formed by items 7^i, 21^i, and 34^i. In BAPQ-SP, secondary dimension (explained variance in first contrast 9.50%, eigenvalue = 1.97, disattenuated correlation with the primary dimension = .49) was formed by items 14, 17, 20, and 32. Lack of unidimensionality impeded to perform further RSM analysis on both Pragmatic Language subscales and Rigid of the BAPQ-EN.

Regarding item separation indices, for the BAPQ-EN, they were between [7.23, 7.64] for Aloof items. For the BAPQ-SP, they were between [7.98, 8.37] for Aloof items and [16.23, 16.94] for Rigid ones.

Correlations of the persons' IRT scores between subscales in BAPQ-SP were: Aloof-Rigidity .63, $p < .001$.

2. Item fit

Infit and outfit mean-squared residual differences for Aloof and Rigid subscales are displayed in Table 1. Regarding the BAPQ-EN Aloof subscale, only Item 18 was slightly upon (1.65) the established boundaries on its outfit index (i.e., [0.6, 1.5]). None item in the BAPQ-SP Aloof and Rigid subscales showed a misfit. Item-total correlations showed adequate values for the BAPQ-EN and BAPQ-SP. The correction of the mistranslation in item 23 resulted in good fit indices and a positive correlation between the item and the subscale (see Table 1) pointing to its adequate functioning similar to what was observed for its English version (BAPQ-EN Aloof subscale).

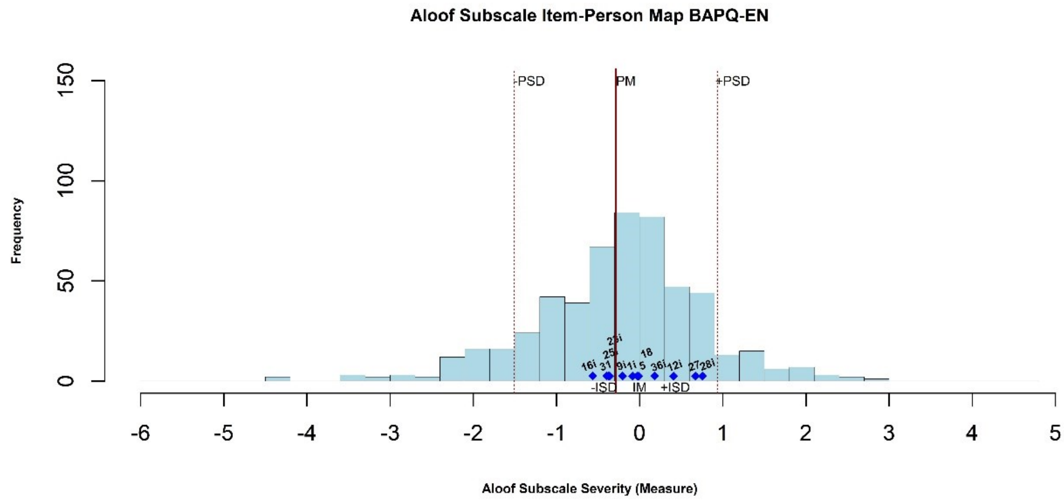
3. Item locations

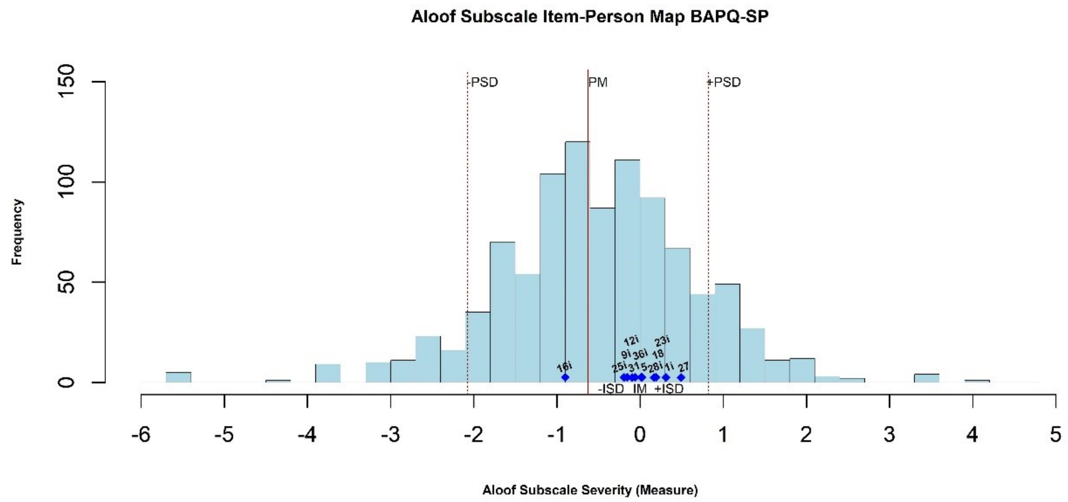
The estimated item and person locations are exhibited in Figure 2 while item locations and estimated parameters are presented in Table 1.

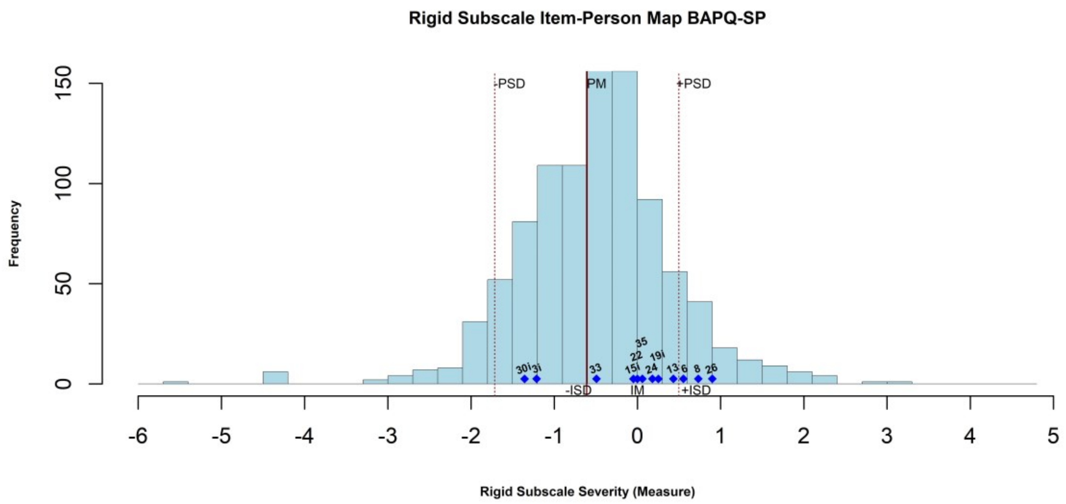
Overall, Aloof item locations showed higher levels of severity (about one item standard deviation above the item mean) than the average Aloof level in the samples for both BAPQ versions. Additionally, some items shared similar locations in the continuum although they refer to different contents (e.g., items 5 and 18 in the BAPQ-EN or items 5,12[i], 31, and 36[i] in the BAPQ-SP). The item severity mean in the Rigid subscale in the BAPQ-SP was about one item standard deviation of the mean of the items. Items 15[i], 22, and 35; items 6 and 13; and items 19[i] and 24 also clustered in the continuum. The remaining five items were spread in different positions below and above the item severity mean.

Figure 2

Person-Item Maps of the BAPQ-EN Aloof, BAPQ-SP Aloof and BAPQ-SP Rigid







Note. Histograms represent the distribution of the person scores regarding each simulated group. Points placed on the x-axis represent the respective item values along the same continuum as person scores. PM = person mean score; +PSD = person standard deviation above the person mean; -PSD = person standard deviation below the person mean; IM = item mean; +ISD = item standard deviation above the item mean; -ISD = item standard deviation below the item mean.

4. Category probability curves

Both Aloof subscales in the BAPQ-EN and Aloof and Rigid in the BAPQ-SP showed adequate category functioning (see Figure 3).

Figure 3

Category Probability Curves of the BAPQ-EN Aloof, BAPQ-SP Aloof and BAPQ-SP Rigid

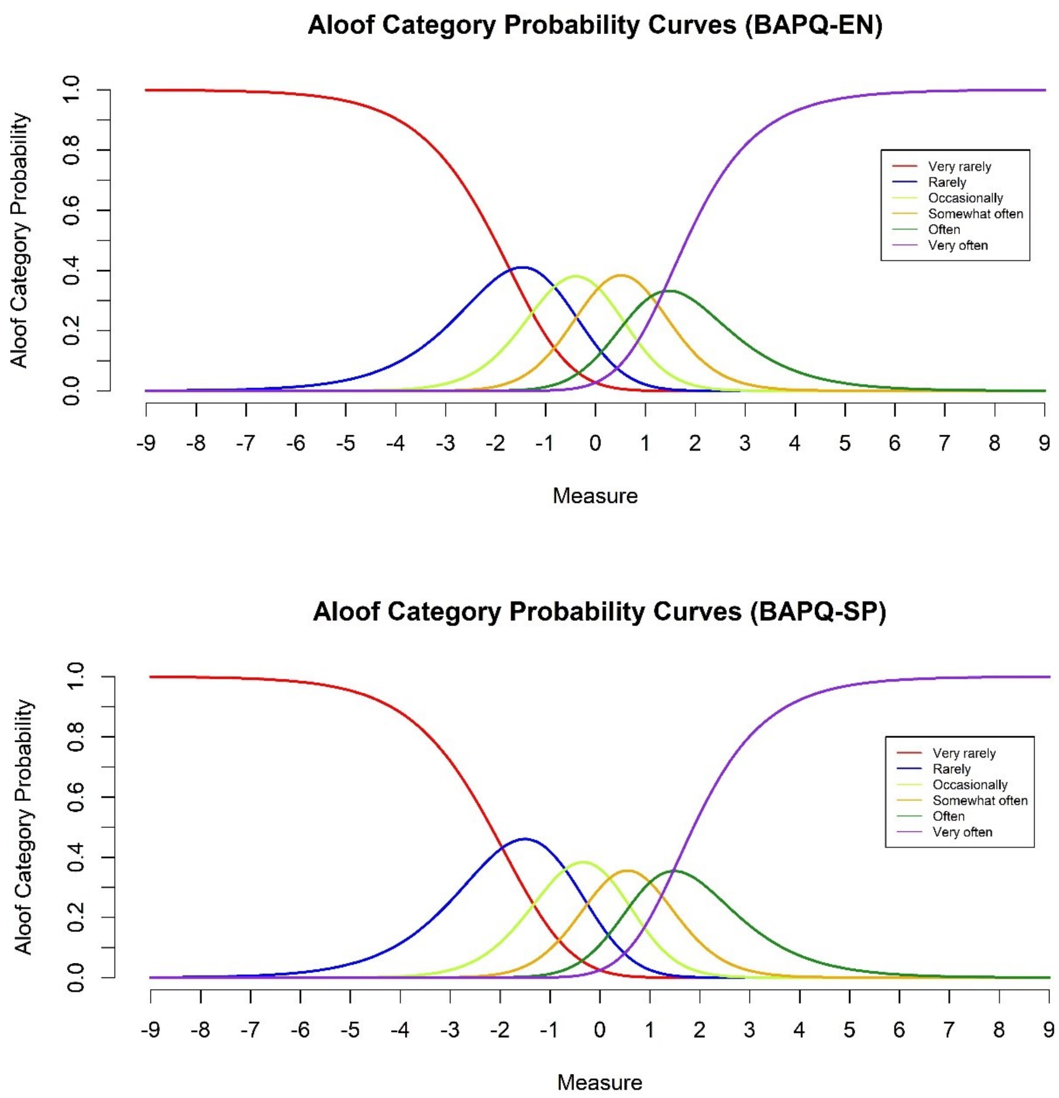
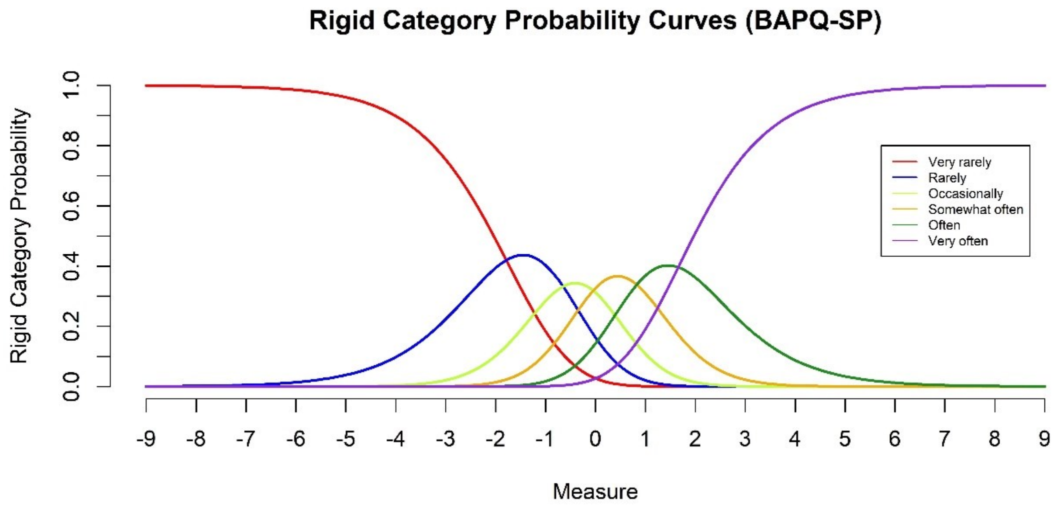


Figure 3

Category Probability Curves of the BAPQ-EN Aloof, BAPQ-SP Aloof and BAPQ-SP Rigid



5. People separation indices and person reliability

The separation index was between [2.60, 3.08] for the BAPQ-EN Aloof subscale ([2.81, 3.29] in the BAPQ-SP, and the scale reliability was between [.87, .90] ([.89, .92] in the BAPQ-SP) that means that the Aloof subscale was able to distinguish almost among three statistically different strata. In the BAPQ-SP Rigid subscale, the separation index was between [2.37, 2.70] and the scale reliability was between [.85, .88], which means that it was able to statistically distinguish between two different strata.

Subscale information functions, as well as standard errors of measurements, can be observed in Figure 4. The point where the Aloof subscale reached its highest precision level corresponded, approximately, with the raw scores 13.14 in BAPQ-EN and 13.22 in BAPQ-SP. The point where Rigid in BAPQ-SP reached its highest precision level corresponded, approximately, with the 13.39 raw score.

Figure 4

Test Information and Standard Error Functions of the BAPQ-EN Aloof, BAPQ-SP Aloof and BAPQ-SP Rigid

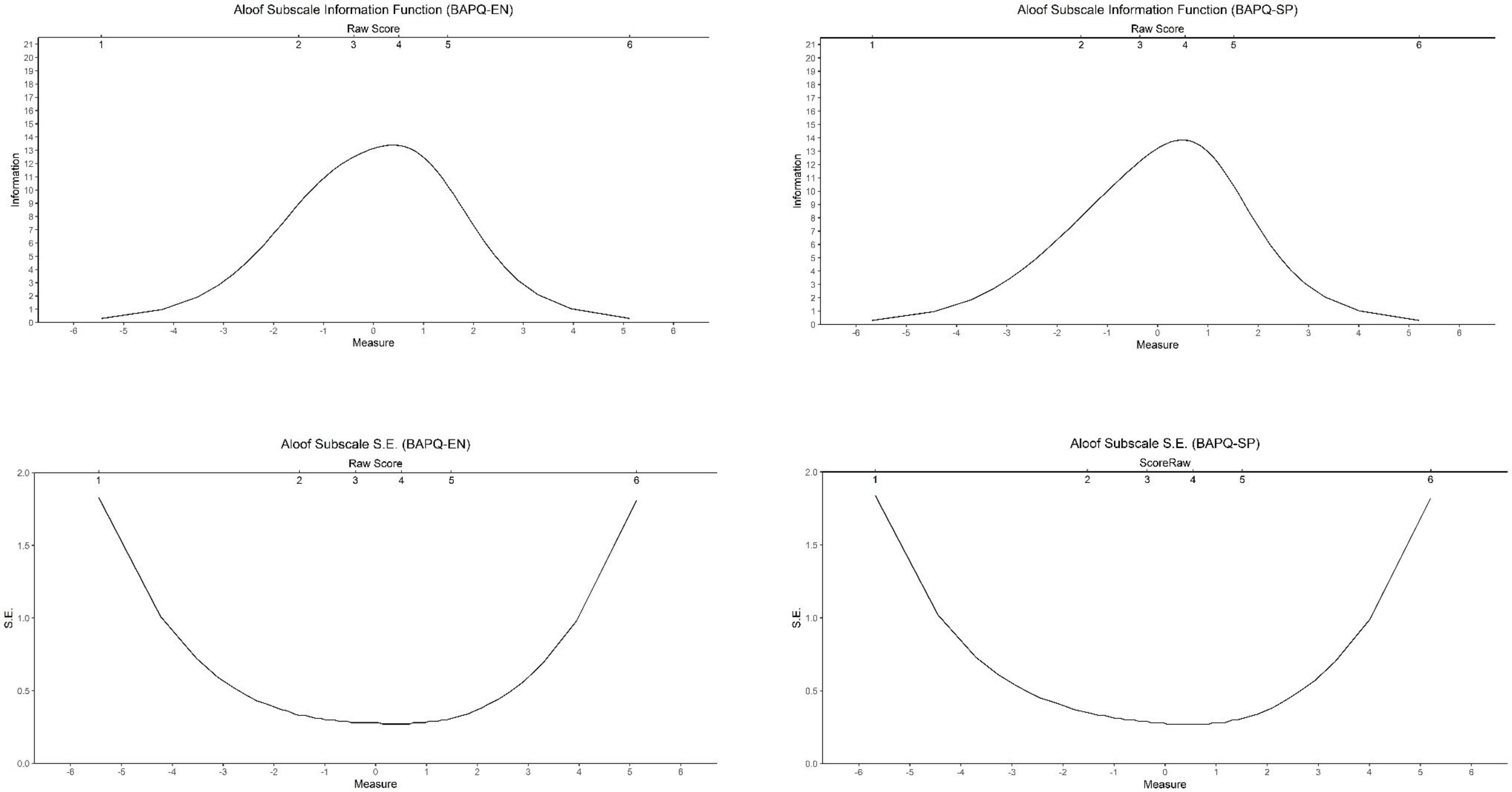
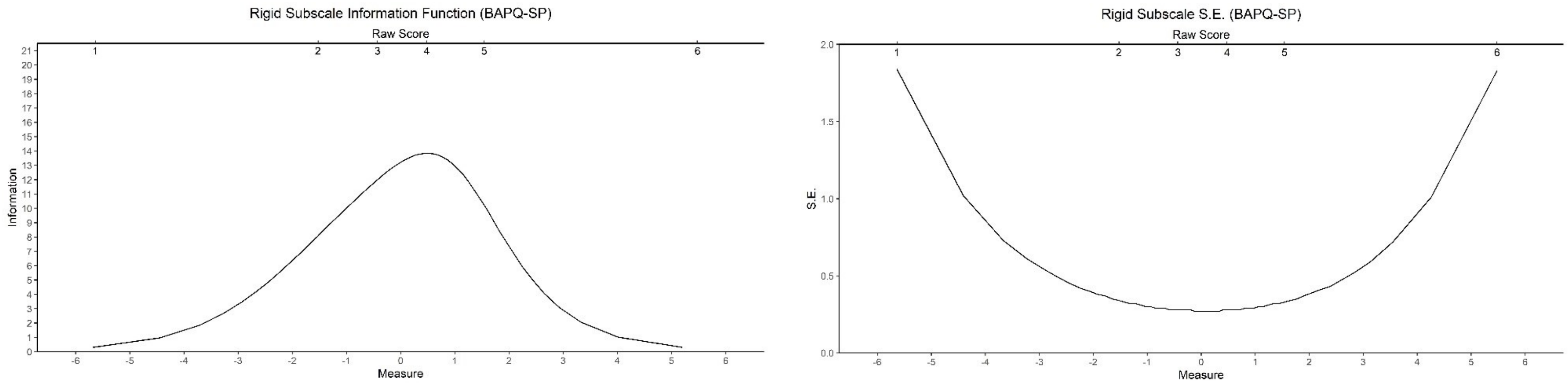


Figure 4

Test Information and Standard Error Functions of the BAPQ-EN Aloof, BAPQ-SP Aloof and BAPQ-SP Rigid



Note. T. I. F. = test information function; S.E. = Standard Errors.

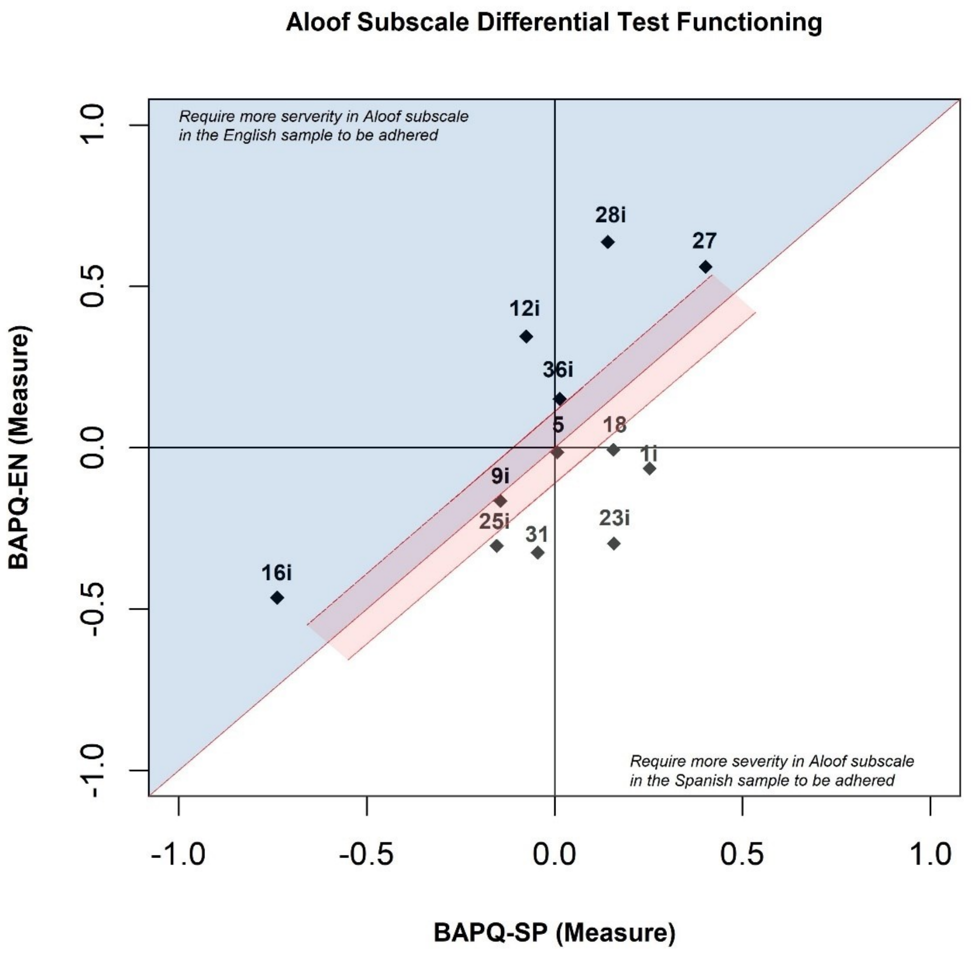
6. Test differential functioning (DFT)

Finally, we checked the Aloof subscale DFT between the English and the Spanish BAPQs.

Ten out of twelve items of the Aloof subscale showed DIF. Only items 5 and 9i had similar functioning. With similar levels of Aloofness, items 1i, 18, 23i, and 25i would be less likely to be endorsed by the Spanish sample than by the English sample, while the opposite happened with items 12i, 16 i, 27, 28i, 31, and 36i.

Figure 5

Identity Plots Of Each Aloof Subscale



Note. Scatterplots represent the severity location of the items scores regarding each Aloof subscale version (BAPQ-EN and BAPQ-SP). Points placed on the x-axis represent the respective item values along the BAPQ-AP severity continuum. Points placed on the y-axis represent the respective item values along the BAPQ-EN severity continuum.

Discussion

In this study, we applied IRT to further study two BAPQ versions (BAPQ-EN and BAPQ-SP; Hambleton et al. 1991). Additionally, we tested the stability of the items' hierarchy and which BAPQ item severities function differentially across versions (Hambleton, 2006).

Significant differences were observed when the same items were applied to both samples (the larger ones are explained further down) being the BAPQ-EN sample the one with overall higher item means. However, only small and medium effect sizes indicated no critical relevance.

Unidimensionality problems in both Pragmatic Language subscales and Rigid one in the BAPQ-EN impeded further rating scale analysis (Bond & Fox, 2015) and, congruently with previous findings, provides negative validity evidence regarding subscales' internal structures. Particularly, Pragmatic Language adverse evidence was reported in Godoy-Giménez et al. (2018), Sasson et al. (2013a), and Sharma et al. (2018); Rigid adverse evidence was reported in Lin et al. (2021), and adverse evidence of both subscales in Stojković et al. (2018). These findings could be better understood under the light of the most updated BAP operationalization (Godoy-Giménez et al., 2021) which revolves around two core domains paralleling it with the last ASD definition (APA, 2013) and includes some variations in the test content.

Regarding item fit, in the BAPQ-EN, we consider that the item in the Aloof subscale (Item 18) that exhibited little misfit could be explained by a lack of concreteness in the item content (any person rating high or low in BAP could agree with it because being polite is independent of the preference for social interactions), yet this slight item misfit could not degrade the measure (Linacre, 2002). By contrast, in the BAPQ-SP, this item did not show any misfit. As this item is located in different parts of the severity continuums of Aloof in both samples, we could suggest that cultural aspects could be playing a role in how different cultures interpret the item. In this regard, if we consider that the item is characterized by a lack of concreteness, it could be possible that the Spanish group understands the item more concretely. However, it would be also reasonable to think that some cultures give more importance to politeness than others. Finally, it is worth mentioning that this item, as other items of the BAPQ, is linked to particularity as in the BAPQ some items are specifically referred to casual interactions with acquaintances (see Godoy-

Giménez et al., 2018; Hurley et al., 2007). Consequently, it could be possible that both samples have differed in how they applied this particularity at the moment to respond to the item. Any other item has not shown misfit nor item 23 from the Aloof subscale in the BAPQ-SP. This result is relevant since, together with a good item-total scale correlation, it supports the alternative translation of this item. In sum, we can conclude that the results of the present study have supported the alternative translations of the Spanish items 4 and 23.

Thirdly, average item severities for both Aloof and BAPQ-SP Rigid subscales, were upper than the average severity level of both community samples. Consequently, subscales' reliability reaches its higher value always for upper scores than both samples mean. This confirms our hypotheses that the items will mainly be situated in middle-upper levels of the continuum indicating that the BAPQ severity levels are more targeted for assessing high levels of BAP, like those observed in parents of ASD children as originally intended in Hurley and cols. (2007). Thus, for those studies focused on community or severity-diverse populations (e.g., Faso et al., 2016; Morrison et al., 2018), it could be advisable to use tests with subtler BAP indicators to enhance the measurement accuracy at lower BAP severity levels.

Regardless of this issue, the items and persons have congruently been scaled according to their severity in both Aloof subscales and the Spanish Rigid one and that implied that the items of the BAPQ subscales are susceptible to being scaled along a severity dimension (severity hierarchies will also be discussed below in terms of validity evidence). Furthermore, our analyses revelated adequate category functioning for both Aloof subscales and the Rigid BAPQ-SP. Thus, even if the test counts on a 6-point scale, we have reported that all the rating options functioned and are most probably chosen at some point in the continuum in both samples. In the same line, adequate reliability and separation indices of both Aloof and the Spanish Rigid subscales inform that subscale hierarchy can statistically differentiate at least among two different strata of BAP severity (i.e., high and low BAP). This represents an important highlight since it goes in line with the ASD severity dimensional approach and opens the door to locate the BAP inside the autism continuum. In this regard, BAP behaviours would not define a qualitatively different group but rather it would comprehend people who could express autistic-like behaviours, at least, at higher or milder severity that makes them more or less functionally independent.

Our second objective was to compare the differential functioning of both Aloof subscales between the Spanish adaptation of the BAPQ (Godoy-Giménez et al., 2018) and its original version (Hurley et al., 2007). Two items showed invariant severity and thus, they would be equally endorsed in both samples by persons with the same severity level of aloofness. By contrast, the lack of invariance in the rest of the items hindered the comparison of the severity hierarchies of both Aloof subscales between the English and Spanish versions.

On the other hand, the location and content of those invariant items need to be discussed. As such, we consider that a mild introverted behaviour could be not to enjoy being in social situations (item 9^{i}: "*I enjoy being in social situations*") which does not imply that the respondent avoids spending time with a few close friends or relatives. By contrast, expressing an instrumental use of the acts of socializing, in general (also involves interactions with relatives, close friends, and partners) could be taken as a more severe indicator of aloofness (item 5: "*I would rather talk to people to get information than to socialize*"). Nevertheless, it is worth mentioning that this last behaviour still locates the person in social interaction and surrounded by others. Coherently, authors have found that ASD parents have shown less interest in purely social interactions, and report having fewer and lower quality friendships (Faso et al., 2016).

Even though the rest of the items have not resulted invariant, both Aloof hierarchies share similar severity order in both groups. This also deserves to be highlighted as it could be taken as additional validity evidence. Specifically, six items of both Aloof hierarchies have shown the same order (from upper to lower locations in the continuum: item 27, 18, 5, 9^{i}, 25^{i}, 16^{i}). Items that shared near placements in both Aloof hierarchies are going to be commented together (upper locations denote more aloofness-like behaviours or more Aloof severity). Item 27 ("*Conversation bores me*") is located in upper locations in both severity continuums and it refers to an open expression of disinterest for establishing social communication in casual interactions with acquaintances. Items expressing less amount of Aloof were 18 ("*When I make conversation it is just to be polite*"), 5 ("*I would rather talk to people to get information than to socialize*"), and 9^{i} ("*I enjoy being in social situations*") which also express this disinterest for social interactions but more subtly and, even when people do not give chance for establishing longer interactions (and do it only for being polite), they do participate in social communication. Finally, items 25[i] and 16[i] are below the mentioned items. Coherently with severity scaling of the items, those two items refer to not having an active pursuit of social situations, social involvement, or social enjoyment, but still, it does not imply being in social situations and participating in them. We could consider that they express milder severity behaviours (e.g., item 25[i] "*I feel like I am really connecting with other people*"; item 16[i] "*I look forward to situations where I can meet new people*").

Importantly, this order in both Aloof hierarchies is also congruent with ASD specifiers in the DSM-V (APA; 2013): Social impaired behaviours at ASD levels could also vary from a decreased interest in social interactions (severity level 1, individual are more functional), to limited initiation of social interactions (severity level 2), and to unusual social approaches restricted to meet needs (severity level 3, individual are less functional; ASD severity specifiers; APA, 2013). This is also an important highlight because both shared Aloof hierarchy orders would be susceptible of being reinserted in the autistic severity continuum (De Groot & Van Strien, 2017; Hoekstra et al., 2008) and also provide validity evidence about the fact that the BAP does not only shares the same autistic behaviours but also that the scalability of the severity of those behaviours follows the same path than in more severe levels of ASD behaviours (APA, 2013).

By contrast, some items were located in a different order in both Aloof hierarchies. For example, items 1[i] and 23[i] are situated below item 27 in upper locations in the BAPQ-SP and item 31[i] in milder locations between items 5 and 9[i] while these items are all located below item 5 in the Aloof BAPQ-EN. In the case of the BAPQ-EN, items 28[i], 12[i], and 36[i] are surrounding item 27 (upper locations of the severity continuum) while in the Aloof BAPQ-SP those items are targeting milder placements. Items with different placements may be referring to varying social situations. In the case of Aloof BAPQ-SP, severe item 1[1] ("*I like being around other people*") and 23[i] *("I am good at making small talk")* refers to social enjoyment and involvement but with the nuance of being surrounded by many people or known people. In the BAPQ-EN, upper items 28[i], 12[i], and 36[i] express being warm and friendly in social interactions (28[i] "*I am warm and friendly in my interactions with others*"; 12[i] "*People find it easy to approach me*"; 36[i] "*I enjoy chatting with people*").

For the item severities that did not remain stable between versions, Hambleton (1994) pointed to "genuine cultural specifics" as a possible explanation of item variance. In this regard, and linked to what we have explained earlier about item 18, we could speculate if this could be suggesting that being polite is an important aspect of social interactions for the English sample while social involvement could be more important in the case of the Spanish one. Nevertheless, these hypotheses should be studied in the future. In the same way, the hierarchical three-level framework that explains how culture affects the perception and diagnosis of psychiatric disorders proposed by Rogler (1993) could be applied to ASD and, in extension, to BAP levels of severity. Accordingly, cultural norms mediate both the endorsement of symptoms and how people rate the severity of symptoms. Therefore, two participants with the same BAP severity but from different countries could express different behaviours responding differently to the BAPQ.

Since the severities of some Aloof items are sensible to the influence of diverse cultural aspects, the direction, as well as the magnitude of these influences, should be established if authors want to further study the interactions among ASD and BAP expressions and the culture. This implies integrating a comprehensive approach taking into account both the invariant and the non-invariant item severities and hierarchies, paying special attention to the influence of the cultures on the development of BAP-related behaviors and preferences but also on their expression either in a natural context or in a test. Ethnically-based cultural norms would modulate the perception of socially undesirable mental symptoms, and, ultimately, the own endorsement of those symptoms (Matson et al., 2017). As such, how people report the behaviours they think they are expressing and their severity may be influenced by the aspects that are considered more worrying within the culture. For example, some studies carried out in the United States have suggested that American parents tend to be more concerned about language delays (Coonrod & Stone, 2004) while Indian parents would tend to have early concerns about social difficulties (Daley, 2004), and Latina mothers may be worried about their child temperament (Ratto et al., 2015).

On the other hand, the severity hierarchy of the Rigid BAPQ-SP should be equally commented. Items sharing near placements of the Rigid continuum are going to be discussed together (upper locations denote more rigidity or more Rigid severity). Items 26 ("*People get frustrated by my unwillingness to bend*") and 8 ("*I have to warm myself up to the idea of visiting an unfamiliar place*") which refers to strong stubbornness or an extreme need for sameness and daily routines were situated above the rest of the items. Below them, we can also find items 13 and 6 that also express a preference for sameness and routines (e.g., 13 "*I feel a strong need for sameness from day to day*", 6 "*People have to talk me into trying something new*") and 24 ("*I act very set in my ways*") and 19[i] (" *I look forward to trying new things*") which also reflect stubbornness, sameness, and routines but more subtly without the nuances of "warm myself" or "unwillingness to bend" that comprehend more severe items (i.e., 26 and 8). Further down in the severity continuum there are items concerning struggling with alternative work procedures (generally about how things must be done) or changes in daily routines (e.g., 35 "*I keep doing things the way I know, even if another way might be better*", 22 "*I have a hard time dealing with changes in my routine*" and 15[i] "*I am flexible about how things should be done*") but, importantly, items do not specify that those difficulties impede to conduct alternative protocols or alternative plans. In lower locations in the continuum, item 33 refers to a preference for fixed work protocol ("*I like to closely follow a routine while working*") but does not express the difficulty component. Finally, following severity order, last placements are taken by items that express a preference for daily routines or not dealing with unexpected plan changes (3[i] "*I am comfortable with unexpected changes in plans*"; 30[i] "*I alter my daily routine by trying something different"*) which we consider behaviours to which almost every person would be comfortable to adhere to. Additionally, we should appreciate that the resulting order of Rigid hierarchy is aligned with the ASD.

Like in the case of Aloof subscales, the resulting order of the Rigid BAPQ-SP hierarchy could open the door to insert behaviours at BAP levels of severity inside the autism continuum (De Groot & Van Strien, 2017; Hoekstra et al., 2008). In this sense, stubbornness, persisting routines and a strong need for sameness (ie., Item 26) could be nearer clinical levels of severity, mostly when inflexible behaviours cause significant interference with individual functioning in one or more contexts

(Item 15[i]; level 1), it comprises difficulty coping with change (Item 8; level 2), and those behaviours marking interfering in all the spheres (level 3; ASD severity specifiers; APA, 2013). Directly linked with this, some autistic-behaviours found upper in the Rigid BAP-SP severity hierarchy are similar to those observed in Obsessive-Compulsive Personality Disorder (e.g., overly rigid and/or stubborn, perfectionism, and very strict work standards; APA, 2013) or even to Obsessive-Compulsive Disorder ones (e.g., obsessional thoughts; APA, 2013). Of importance, obsessive-compulsive behaviours have been reported comorbid to autism spectrum disorder (e.g., Meier et al., 2015; Micali et al., 2004) and this would also support the resulting Rigid severity order and the reinsertion of this dimension inside the autism continuum.

Altogether, the presented results lead us to think that the BAPQ could have been a useful measurement tool for measuring the BAP in English populations with middle-high BAP severity levels according to the original BAP operationalization (Piven et al., 1997a). Nevertheless, this study, in conjunction with others, has provided adverse evidence about the internal structure of the BAPQ (e.g., Godoy-Giménez et al., 2018; Lin et al., 2021; Sasson et al., 2013a; Sharma et al., 2018; Stojković et al., 2018). These findings could be better understood considering the most updated BAP operationalization (Morrison et al., 2018; Sasson et al., 2013b; Godoy-Giménez et al., 2021) and suggest that a new BAP test based on an updated BAP definition (as the middle expression of the current two-dimensions operationalization of the ASD; APA, 2013; see Godoy-Giménez et al., 2021) should be built. The new test would aim to measure BAP covering all autism subthreshold levels of expression including those in the general population. Additionally, to accomplish the requirements of the current international research context we believe that this new test should aim to inform about how the severity of the items could vary among cultures.

Declarations

This work was supported by grant Grants PID2019-110066GB-I00 funded by MCIN/AEI/ 10.13039/501100011033 to AFE; and by the Spanish Ministry of Education, Culture, and Sport to MGG (FPU2015-015). Authors have no known conflict of interest to disclose.

Compliance with Ethical Standards

This study received ethics approval from the local Human Research Ethics Committees of both universities. The recruitment process was conducted under the approved guidelines and the Declaration of Helsinki. Written informed consent was obtained from all participants in advance. Participants in both samples received two-course credits for their collaboration

References

Álvarez-Couto, M., García-Villamisar, D., Sáez-Suanes, G. P., & Roquete, M. D. O. (2021). Variables predicting the association between autistic traits and externalizing symptoms among young adults. *Advances in Autism*. Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/AIA-06-2020-0036

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

Bai, D., Yip, B. H. K., Windham, G. C., Sourander, A., Francis, R., Yoffe, R., ... & Sandin, S. (2019). Association of genetic and environmental factors with autism in a 5-country cohort. JAMA Psychiatry, 76(10), 1035-1043. https://doi.org/10.1001/jamapsychiatry.2019.1411

Bang, P., Strömberg, M., Meera, S. S., & Igelstrom, K. (2021). Brief Report: The Broad Autism Phenotype in Swedish Parents of Children with and without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*. 1-8. https://doi.org/10.1007/s10803-021-05302-3

Berry, J. W., Poortinga, Y. H., Segall, H. H., & Dasen, P. R. (1992). Cross-cultural psychology: research and applications. Cambridge: Cambridge University Press.

Bölte, S., Poustka, F., & Constantino, J. N. (2008). Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Research*, *1*(6), 354-363. https://doi.org/10.1002/aur.49

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). L. Erlbaum.

Broderick, N., Wade, J. L., Meyer, J. P., Hull, M., & Reeve, R. E. (2015). Model Invariance Across Genders of the Broad Autism Phenotype Questionnaire. *Journal of autism and developmental disorders, 45*, 3133-3147. https://doi.org/10.1007/s10803-015-2472-z

Camodeca, A., & Voelker, S. (2016). Automatic and controlled processing and the Broad Autism Phenotype. *Psychiatry Research*, *235*, 169–176. https://doi.org/10.1016/j.psychres.2015.11.012

De Groot, K., & Van Strien, J. W. (2017). Evidence for a broad autism phenotype. *Advances in Neurodevelopmental Disorders*, *1*(3), 129-140. https://doi.org/10.1007/s41252-017-0021-9

Faso, D. J., Corretti, C. A., Ackerman, R. A., & Sasson, N. J. (2016). The broad autism phenotype predicts relationship outcomes in newly formed college roommates. *Autism, 20*(4), 412-424. https://doi.org/10.1177/1362361315585733

Godoy-Giménez, M., González-Rodríguez, A., Cañadas, F., Estévez, A. F., & Sayans-Jiménez, P. (2018). Psychometric Properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements. *Journal of autism and developmental disorders, 48*(3), 770-78. https://doi.org/10.1007/s10803-017-3438-0

Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment, 10*(3), 229-244.

Hambleton, R. K. (2006). Good Practices for Identifying Differential Item Functioning. *Medical Care, 44*(11), S182-S188. Retrieved from http://www.jstor.org/stable/41219518

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Sage Publications.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Sage publications.

Hurley, R. S., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of autism and developmental disorders, 37*, 1679-1690. https://doi.org/10.1007/s10803-006-0299-3

Ingersoll, B. & Wainer, A. 2014. The Broader Autism Phenotype. In F.R. Volkmar, R. Paul, S.J. Rogers, K.A. Pelphrey (Eds). *Handbook of Autism and Pervasive Developmental Disorders* (4th ed.). John Wiley & Sons. https://doi.org/10.1002/9781118911389

Jakobson, Lorna S.; Pearson, Pauline M.; Kozub, Zofia; Hare, Colleen; Rigby, Sarah N. (2018). Links between traits associated with the broad autism phenotype and empathy and young adults' ability to decode speaker intentionality.

Research in Autism Spectrum Disorders, 50, 11–21. https://doi.org/10.1016/j.rasd.2018.03.001

Jamil, Ruby; Gragg, Marcia N.; DePape, Anne-Marie (2017). The broad autism phenotype: Implications for empathy and friendships in emerging adults. *Personality and Individual Differences, 111,* 199–204. https://doi.org/10.1016/j.paid.2017.02.020

Kim, H. U. (2012). Autism across cultures: Rethinking autism. *Disability & Society, 27*(4), 535-545. https://doi.org/10.1080/09687599.2012.659463

Landa, R., Piven, J., Wzorek, M., Gayle, J., Chase, G., & Folstein, S. (1992). Social language use in parents of autistic individuals. *Psychological Medicine, 22*, 245-254. https://doi.org/10.1017/S0033291700032918

Landry, O., & Chouinard, P. A. (2016). Why we should study the broader autism phenotype in typically developing populations. *Journal of Cognition and Development, 17*, 584-595. https://doi.org/10.1080/15248372.2016.1200046

Lin, J., Fu, X., Dai, S., Zeng, Y., Yang, R., Liu, R., ... & Ou, J. (2021). The Broad Autism Phenotype Questionnaire (BAPQ): Strengths, weaknesses and future improvements in Chinese version. *Asian Journal of Psychiatry, 60*, 102655. https://doi.org/10.1016/j.ajp.2021.102655

Linacre, J. M. (2018). *Winsteps® Rasch measurement computer program. User's Guide.* Beaverton, or: Winsteps.com.

Masters, G. N., & Wright, B. D. (1984). The essential process in a family of measurement models. *Psychometrika*, *49*(4), 529–544. https://doi.org/10.1007/bf02302590

Matson, J. L., Matheis, M., Burns, C. O., Esposito, G., Venuti, P., Pisula, E., Misiak, A., Kalyva, E., Tsakiris, V., Kamio, Y., Ishitobi, M., & Goldin, R. L. (2017). Examining cross-cultural differences in autism spectrum disorder: A multinational comparison from Greece, Italy, Japan, Poland, and the United States. *European psychiatry : the journal of the Association of European Psychiatrists*, *42*, 70–76. https://doi.org/10.1016/j.eurpsy.2016.10.007

Meera, S. S., Girimaji, S. C., Seshadri, S. P., Philip, M., Shivashankar, N., Morgan, P., & Piven, J. (2015). Translation of the Broad Autism Phenotype Questionnaire to an Indian language: A description of the process. *Asian Journal of Psychiatry, 15*, 62–67. https://doi.org/10.1016/j.ajp.2015.04.013

Meier, S. M., Petersen, L., Schendel, D. E., Mattheisen, M., Mortensen, P. B., & Mors, O. (2015). Obsessive-Compulsive Disorder and Autism Spectrum Disorders: Longitudinal and Offspring Risk. *PloS one*, *10*(11), e0141703. https://doi.org/10.1371/journal.pone.0141703

Micali, N.; Chakrabarti, S.; Fombonne, E. (2004). *The Broad Autism Phenotype: Findings from an Epidemiological Survey. Autism, 8, 21-7.* https://doi.org/10.1177/1362361304040636

Morrison, K. E., Chambers, L. K., Faso, D. J., & Sasson, N. J. (2018). The content and function of interests in the broad autism phenotype. *Research in Autism Spectrum Disorders*, *49*, 25-33. https://doi.org/10.1016/j.rasd.2018.02.002

Neil W. Scott; Peter M. Fayers; Neil K. Aaronson; Andrew Bottomley; Alexander de Graeff; Mogens Groenvold; Chad Gundy; Michael Koller; Morten A. Petersen; Mirjam A.G. Sprangers (2009). *A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. , 62(3), 0–295. https://doi.org/*10.1016/j.jclinepi.2008.06.003

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997a). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry, 154*(2), 185–190. https://doi.org/10.1176/ajp.154.2.185

Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997b). Personality and language characteristics in parents from multiple-incidence autism families. *American Journal of Medical Genetics, 74*(4), 398–411. https://doi.org/10.1002/(SICI)1096-8628(19970725)74:4<398::AID-AJMG11>3.0.CO;2-D

Rogler, L. H. (1993). Culturally sensitizing psychiatric diagnosis: A framework for research. *Journal of Nervous and Mental Disease, 181*(7), 401-408. http://dx.doi.org/10.1097/00005053-199307000-00001

Sasson, N. J., Faso, D. J., Parlier, M., Daniels, J. L., & Piven, J. (2014). When father doesn't know best: Selective disagreement between Self-Report and informant report of the broad autism phenotype in parents of a child with autism. *Autism Research, 7*(6), 731-739. https://doi.org/10.1002/aur.1425

Sasson, N. J., Lam, K. S. L., Childress, D., Parlier, M., Daniels, J. L., & Piven, J. (2013a). The Broad Autism Phenotype Questionnaire: Prevalence and Diagnostic Classification. *Autism Research, 6*, 134–143. https://doi.org/10.1002/aur.1272

Sasson, N. J., Lam, K. S., Parlier, M., Daniels, J. L., & Piven, J. (2013b). Autism and the broad autism phenotype: familial patterns and intergenerational transmission. *Journal of Neurodevelopmental Disorders, 5*, 11. https://doi.org/10.1186/1866-1955-5-11

Seidman, I., Yirmiya, N., Milshtein, S., Ebstein, R. P., & Levi, S. (2012). The Broad Autism Phenotype Questionnaire: mothers versus fathers of children with an autism spectrum disorder. *Journal of autism and developmental disorders, 42*(5), 837-846. https://doi.org/10.1007/s10803-011-1315-9

Sharma, Ujjwal; Bhushan, Braj (2018). The Hindi Adaptation and Standardization of BAPQ: A Brief Tool for Assessing Broad Autism Phenotype. *Asian Journal of Psychiatry*, S1876201818307494–. https://doi.org/10.1016/j.ajp.2018.10.008

Shi, L.-J., Ou, J.-J., Gong, J.-B., Wang, S.-H., Zhou, Y.-Y., Zhu, F.-R., … Luo, X.-R. (2015). Broad autism phenotype features of Chinese parents with autistic children and their associations with severity of social impairment in probands. *BMC Psychiatry*, *15*(1). https://doi.org/10.1186/s12888-015-0568-9

Stojković, I., Dučić, B., Kaljača, S., & Đorđević, M. (2018). Construction and psychometric evaluation of a short form of the Broad Autism Phenotype Questionnaire. *Psihologija, 51*(2), 243-258. https://doi.org/10.2298/PSI170619014S

Sucksmith, E., Roth, I., & Hoekstra, R. A. (2011). Autistic traits below the clinical threshold: re-examining the broader autism phenotype in the 21st century. *Neuropsychology review, 21*, 360-389. https://doi.org/10.1007/s11065-011-9183-9

Wainer, A. L., Ingersoll, B. R., & Hopwood, C. J. (2011). The Structure and Nature of the Broader Autism Phenotype in a Non-Clinical Sample. *Journal of Psychopathology and Behavioral Assessment, 33*, 459–469. https://doi.org/10.1007/s10862-011-9259-0

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8,* 370-371 retrieve from https://www.rasch.org/rmt/contents.htm

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis.* MESA press.

CHAPTER FIVE.a. THE BASELINE OF THE STUDY 3:

WHAT DO EXPERTS HAVE TO SAY ABOUT THE TWO-DOMAIN OPERATIONALIZATION OF THE BAP? THEORETICAL EVIDENCE ABOUT ITS NEW DIMENSIONAL STRUCTURE

Due to limitations of lengths specified in the guidelines of the Journal of Autism and Developmental Disorders (journal where Study 3 has been published), some the results of the first step of the Study 3 were removed from the manuscript. In this Chapter Five.a., I present the results of the baseline of Study 3. One of the tasks that the experts had to do was to revise and rate the adequateness of the BAP operational definition. As I have explained in Chapters One and Two, my proposal has been an extension of the ASD dimensional structure to grasp BAP levels of severity. Besides expressing their degree of agreement with this BAP operationalization, experts also rated the relevance of each specific item for the subdomain it has been assigned to and how well the items represented the subdomain (see Study 3.b.).

Seventeen of the twenty experts agreed that the BAP operationalization was well-represented by the seven subdomains included in the BAP operationalization (85%) while the remaining three experts considered it very well-represented (15%). None of the experts believed that the construct needed an alternative operationalization (it was bad represented). One of the experts also suggested that the construct of the BAP could be enriched by the inclusion of contents related to information processing and different cognitive profiles particularly regarding perceptive processing. Regarding how well the items represented their specific subdomain, as it is deeply exposed in the Chapter Five.b., subdomain, 19 of the experts (95%) considered that the BAP construct was well-represented by the items of the three BAP tools (e.g., AQ, BAP, and SRS-II) whilst one response was missing (5%)

CHAPTER FIVE.b. STUDY 3:

IS IT POSSIBLE TO ASSESS THE TWO-DOMAIN DEFINITION OF THE BROAD AUTISM PHENOTYPE USING THE AVAILABLE MEASUREMENT TOOLS?

Manuscript published as:

Godoy-Giménez, M., González-Rodríguez, A., Cañadas, F., Estévez, A. F., & Sayans-Jiménez, P. (2021). Is it possible to assess the two-domain definition of the Broad Autism Phenotype using the available measurement tools? *Journal of autism and developmental disorders.* https://doi.org/10.1007/s10803-021-05158-7

Is it Possible to Assess the Two-Domain Definition of the Broad Autism Phenotype Using the Available Measurement Tools?

Godoy-Giménez, M[1]., González-Rodríguez, A[1]., Cañadas, F[1,2]., Estévez, A. F.[1,2]., and Sayans-Jiménez, P[1].

[1]*Department of Psychology, University of Almeria, 04120 Almería, Spain*

[2]*CERNEP Research Centre, University of Almeria, 04120 Almería, Spain*

Author Note

Marta Godoy-Giménez https://orcid.org/0000-0002-2298-0113

Antonio González-Rodríguez https://orcid.org/0000-0001-6432-829X

Fernando Cañadas http://orcid.org/0000-0001-7689-0132

Angeles F. Estévez http://orcid.org/0000-0003-0067-0840

Pablo Sayans-Jiménez https://orcid.org/0000-0001-6681-5729

Correspondence concerning this article should be addressed to Pablo Sayans-Jiménez, Department of Psychology, University of Almería, 04120 Almería, Spain. Email: psj242@ual.es, and Angeles F. Estévez, Department of Psychology, Cernep Research Centre, University of Almería, 04120 Almería, Spain. Email: mafernan@ual.es.

Abstract

Although, the operationalization of the autism spectrum disorder has been updated around two domains, the broad autism phenotype (BAP) one has not. Additionally, the items of the three common BAP measures, the Broad Autism Phenotype Questionnaire (BAPQ), the Autism Quotient, and the Social Responsiveness Scale (SRS), remain organized around a non-consensual number of factors. We explored whether the items of these measures matched with the two-domain operationalization through a parallel analysis, which has suggested two main components, and two expert judgments which have assessed item wording, relevance, and construct representativeness. A remaining pool of 48 BAP-relevant items suggested a possible under-representation of two subdomains. Despite the relevance of all the BAPQ items, only the SRS ones tapped in all subdomains.

Keywords: broad autism phenotype, test content, expert judgment, BAPQ, SRS, AQ

Is it Possible to Assess the Two-Domain Definition of the Broad Autism Phenotype Using the Available Measurement Tools?

The broad autism phenotype (BAP) has been described as a set of subthreshold features qualitatively similar to those existing in autism spectrum disorder (ASD), which is continuously distributed and spreads beyond ASD family members into the general population (Constantino & Todd, 2003; Hoekstra et al., 2007; Hurst et al., 2007; Stewart & Austin, 2009). As pointed out by numerous authors (e.g., Morrison et al., 2018; Wainer et al., 2011), BAP traits correspond directly to the primary characteristics of ASD: Social Communication and Social Relation Impairment (SCI) alongside a Pattern of Restricted Repetitive Behaviours and Interests (RRBs; American Psychiatric Association [APA], 2013). Even though the definition of ASD has been updated over time, the most popular BAP definition has barely changed and has not been revised according to the most recently updated definition of ASD (APA, 2013). This has generated a discrepancy between ASD and BAP conceptualizations that, in our opinion, needs to be addressed.

The reharmonization of the BAP and ASD constructs will highly contribute to the field as it will operationalize the BAP in the spectrum, increasing its understanding and, providing further information regarding how core autistic deficits are expressed differentially on each severity level. Likewise, it will also help to uncover autistic genetic mechanisms (Gaugler et al., 2014; Robinson et al., 2011) by exploring the presence of BAP in parents of children diagnosed with ASD. In the same way, it will also increase the understanding of the developmental path of autism in elder adults (Stewart et al., 2018) which has not been very studied yet. Finally, we can as well take advantage of BAP samples to test assessment and intervention procedures before applying them to clinical and more sensible samples.

On the other hand, accepting that both constructs essentially represent the same spectrum of traits to a different degree, could bring many advantages for several populations. As such, uncertain cases of those individuals who do not have a clear ASD diagnosis but have shown highly impaired BAP behaviours and, in lower intensity, independent and functional adults with some core autistic behaviours that impoverish their social interactions, as well as those capable of camouflaging or smoothing specific deficits (developing ways to cope with them; Hull et al., 2017; Livingston et al., 2019; Mandy, 2019), could also benefit from well-established autism assessment protocols and interventions which will represent, indeed, a clinical milestone.

In this regard, the reconciliation of BAP and ASD will drive certainly to restore any previous measurement process. According to the Standards (2014), whether two tests have defined the same construct differently and, thus, contain different internal structures, those tests are assessing essentially two different things. The BAP, as the autistic phenotype, has been proposed to share ASD internal structure, and, consequently, tests that aim to measure the BAP should be developed upon an updated two-factor internal structure. Conducting studies to establish conclusions on ASD populations through BAP samples necessarily imply the existence of a measurement tool capable of measuring an updated BAP operationalization. For all the abovementioned reasons, in this work, we aimed to confirm whether it is possible to integrate the items of the most frequently used tests for assessing the BAP to provide a BAP measure that would be aligned with the current operationalization of ASD.

The Outdated Operational Definition of the BAP

The BAP was first reported by Leo Kanner (1943) who observed that the parents of children diagnosed with ASD presented subtle expressions of autistic-like traits such as an obsession with details, social awkwardness, and rigid behaviours. Later, more formal investigations conducted by Piven et al., (1997a, 1997b) identified autistic-like core deficits in first-degree relatives of people diagnosed with ASD. Those deficits established the foundations for the first operational definition of BAP, which was developed by Hurley et al. (2007), who, after reviewing outcomes of the studies conducted in the previous two decades, defined the BAP as a set of subclinical personality characteristics and language deficits clustered around three main domains; aloof personality, rigid personality, and pragmatic language impairment paralleling its definition with that proposed for ASD by the DSM-IV-TR (American Psychiatric Association [APA], 2000).

According to the literature, an aloof personality involved a lack of social responsiveness (Constantino et al., 2006), reduced social abilities (Wheelwright et al., 2010), and scarce social engagement (Whitehouse et al., 2010). Likewise, a rigid personality was manifest in behavioural rigidity, a tendency toward perfectionism, stubbornness, and stereotyped behaviours (e.g., Losh et al., 2008; Murphy et al., 2000; Narayan et al., 1990). Finally, pragmatic language deficits were related to supra-linguistic aspects such as problems in respecting turn-taking in speaking, becoming side-tracked in conversations, and difficulties in maintaining the topic of conversations (Seidman et al., 2012).

Although this is now a more mature field, there is still no universal agreed operationalization of the BAP, but rather a quantitative and qualitative amalgam of traits that vary according to the measurement method used to assess the phenotype (Wainer et al., 2011). Those varied features have been, sometimes, difficult to cluster so as to conform to a consensual structure for the BAP enabling to reinsert it inside the autism continuum. In light of more recent research, it makes no sense that the phenotypic expression of ASD, which had emerged for improving the knowledge about its aetiology, would diverge from its mother disorder and, thus, consensus could be reached in the conceptualization of the BAP by aligning it with the updated definition of the ASD (APA, 2013). According to this definition, ASD has been defined as a continuum of increasing severity. Some authors have previously argued in favour of this idea, associating the lesser, non-clinical expressions of autism, with the BAP (Bolton et al., 1994; Constantino & Todd, 2003; Piven & Palmer, 1999). In this regard, some studies have already proposed that the BAP should comprise only two characteristic traits, these being conceptualized as both social and non-social expressions of the BAP, where social traits have included both social impairment in social relations and in social communication; whilst non-social traits have constituted a rigid personality (Morrison et al., 2018; Sasson et al., 2013b).

The Measurement of the BAP

Although the BAP has traditionally been assessed through structured and extensive interviews designed to evaluate personality (i.e., M-PAS-R; Piven et al., 1994), the use of brief psychometric self or informant-reports has increased, reducing time costs and enhancing objectivity. Among these questionnaires, the Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001), the Social Responsiveness Scale (SRS; Constantino & Gruber, 2005), and the Broad Autism

Phenotype Questionnaire (BAPQ; Hurley et al., 2007) have been the most widely used (for a review see, Ingersoll & Wainer, 2014).

Despite having plenty of benefits, researchers and clinicians should be particularly careful regarding the theoretical and empirical evidence supporting the interpretation of these test scores. In particular, following the changes in the operationalization of ASD from three to two general domains (APA, 2013), it might be advisable to draw special attention to different sources of validity evidence such as the internal structure or test content. Adverse evidence would imply the need to question the interpretation of the test scores (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 2014). As stated above, there has been a lack of correspondence between the operationalization of the BAP (understood as the subclinical expression of autism which includes both impairment in social interactions and rigid behaviours) and the content of the tests most frequently used for its evaluation. Of equal importance is the fact that this lack of correspondence could also be observed between the operationalization of the construct and the internal structure of these tests.

The three aforementioned measures have reflected this lack of correspondence. Thus, the AQ—which was originally designed for assessing autism in adults with typical range IQs—comprises five content-domains and five factors: Social Skills, Attention Switching, Attention to Detail, Communication, and Imagination (Baron-Cohen et al., 2001). In spite of contradictory data in the literature, some studies have supported the clustering of the five subscales into a three-factor model: Social skills, Details/patterns, and Communication/mindreading (English et al., 2019; Hurst et al., 2007; Russell-Smith et al., 2011). Similarly, the second adult version of the SRS (SRS-2) measures autism in adults with typical development and includes five subscales: Social Awareness, Social Cognition, Social Communication, Social Motivation, and Restrictive and Repetitive Behaviors (Bruni, 2014; Constantino & Gruber, 2012). It includes two subscales compatible with the DSM-5: Social Communication and Interaction, and Restricted Interests and Repetitive Behaviours. Scores on these subscales have facilitated the comparison of symptoms with DSM-5 diagnostic criteria for ASD. These comparisons could help to determine whether a person meets the most current diagnostic criteria for ASD (Constantino & Gruber, 2012). As support for this idea, Frazier et al. (2014) observed that SRS-2 subscales could be encompassed by the following two-structured factors: Social Communication/Interaction (SCI), and Restricted/ Repetitive Behaviour (RRB; Frazier et al., 2014). Finally, the BAPQ, the tool originally designed for assessing the BAP in parents of children diagnosed with ASD, includes three factors: Aloofness, Pragmatic Language Deficits, and Rigidity (for a review, see Hurley et al., 2007; Sasson et al., 2013a).

Despite the diverse specifications of these tests, some authors have suggested that these all could serve to assess the same underlying BAP structure. In an attempt to search for the similarities between these tests and to obtain an empirically-based latent structure of the BAP in a nonclinical sample, Wainer et al. (2011) conducted a conjoint exploratory factor analysis by collapsing the AQ, the SRS-A (adult self-report version; Constantino & Gruber, 2005), and the BAPQ subscales. They concluded that the subscales could conform to a three-factor structure similar to that proposed for ASD by the DSM-IV-TR (APA, 2000). These three factors, named Aloof Personality, Pragmatic Language Difficulties, and Rigid Personality, matched with the BAPQ factors and with one of the most accepted descriptions and operationalizations of the BAP (Piven et al., 1997a, 1997b).

In view of their findings, Wainer et al. (2011) also discussed their resulting BAP structure and suggested that more research was needed to clarify whether the core areas of impairment observed in ASD (and, by extension, the BAP as its non-clinical expression) could be grouped into two core domains. This was in line with recent literature that has suggested the need for a new conceptualization of the phenotype aligned with the current definition of ASD (APA, 2013), centred around two core domains: social and non-social areas (Morrison et al., 2018; Sasson et al., 2013b).

To take the work of Wainer et al. (2011) one step further, the present study aimed to extend and revise their findings by exploring the connection between the AQ, the BAPQ, and the SRS-2 subscales with an updated conceptualization of BAP according to two main dimensions or domains: Deficits in Social Communication and Social Interaction (hereafter, SCI BAP) and Restricted, Repetitive Patterns of Behaviours, Interests, or Activities (hereafter, RRB BAP). To this end, our study firstly explored the internal structure resulting from applying parallel analysis

(Horn 1965) to the set of total scores of each subscale of the mentioned tests (first objective, parallel analysis). Secondly, we took the first step towards bridging the gap between the operationalization and measurement of the BAP by selecting the most relevant items for its measurement. For this second objective, a group of collaborators from our research lab allocated the items of these questionnaires, according to their content, to the seven ASD subdomains (Loevinger 1957). Furthermore, they verified their itemmetric properties (see Angleitner et al. 1986; Grant & Davis 1997).

Finally, we recollected quantitative information regarding content validity (Lynn, 1986). For achieving our third objective, a broad group of experts evaluated the relevance (Ebel & Frisbie, 1972; Haynes et al., 1995) and the representativeness (Haynes et al., 1995; Lynn, 1986) of the selected items for assessing BAP features based on the updated operationalization aligned with the current definition of ASD (Armstrong et al., 2005; Beck & Gable, 2001; Delgado-Rico et al., 2012.

Following the results presented in Wainer et al. (2011), we hypothesized, that the subscales of the AQ, the BAPQ, and the SRS-2 could be clustered into three components of variance that could correspond to the original conceptualization of BAP (first objective; Hurley et al., 2007; Piven, et al., 1997a, 1997b). Further, we expected to find some problems regarding the formal aspects of items (second objective; see Angleitner et al., 1986). For instance, some items could present high levels of social desirability, low levels of self-reference, lack of concreteness or be difficult to understand. In addition, since items stem from different tests with different operationalizations of the BAP (in the case of BAPQ) and ASD (in the case of AQ and SRS-2), we hypothesized that their relevance for assessing one of the seven BAP subdomains could equally affect their final selection (third objective). For example, Item 11 SRS-2 states “I have got self-confidence” which could not be relevant since it did not capture any of the key contents of the seven BAP subdomains. Finally, since the items were constructed within the framework of previous BAP or ASD definitions, we anticipated that some of the seven BAP subdomains could be under-represented (third objective).

Method

Participants

Parallel Analysis

A total of 349 undergraduates [26.1% men, M (SD) = 21.56 (4.63); 73.9% women, M (SD) = 20.88 (4.01)] completed a booklet which contained the questionnaires. Participants were recruited from the Degree of Psychology of the University of Almeria through incidental and snowball sampling. Specific data on socioeconomic status and ethnicity were not recorded. None of the students reported severe or genetic conditions or a history of psychiatric disorders.

Item Selection and Assignment According to their Content and Itemmetric Properties

Five collaborators of our research lab took part in this phase of the study (FC, AFE, MGG, and AGR were experts in neuropsychology and PSJ in psychometrics). They were purposely selected due to their expertise in the core fields associated with our intended objectives (i.e., autism-related traits and test construction).

Expert Judgment

Twenty judges (18 professionals and two academic experts) participated in this second phase. The criteria for selecting them were, in the case of professionals, to have been working with ASD people and their families during the past five years. Academic experts must have completed a doctoral dissertation and have more than five published articles on the subject. Moreover, to ensure the representativeness of the sample of experts (Davis, 1992), we developed a grid with fields related to ASD, in both professional and academic areas (see Table 1). Following the authors' recommendations (Gable & Wolf, 1993; Tilden et al., 1990) and due to the large pool of items, an initial sample of 32 experts (24 professional experts and eight academic experts) were selected. After making contact with them, 23 agreed to participate in the study. Three of them were subsequently deleted from the data due to missing information in their booklets. Regarding formal education of final grid of experts, eleven of them had studied a Degree in Psychology, two have a Degree in Speech Therapy, four had a Degree in Special Education, one did a Degree in Pedagogy, and two had higher studies in Neuropaediatry.

Table 1

Grid of ASD Related Areas of Expertise Covered in Expert Judgment

Professional experts (n= 18)					
Health field (n = 11)		Social field (n = 2)		Educational field (n = 5)	
Early care		Professional workers in autism associations	n = 2	School counsellors	n = 2
Speech therapist	n = 2				
Child neuropsychologist	n = 2				
Child psychologist	n = 2				
Neuropaediatric Unit	n = 1			Workers in therapeutic pedagogy	n = 3
Child psychologist (psychologist specialist in clinical psychology)	n = 4				
Academic experts (n =2)					
Expert in educational and developmental psychology (n = 1)					
Expert in clinical child psychology (n = 1)					

Instruments

Parallel Analysis

The Broad Autism Phenotype Questionnaire (Spanish Self-report Version). The BAPQ-SP (Godoy-Giménez et al., 2018; Hurley et al., 2007) was a 36-item screening questionnaire specifically designed to assess the BAP in relatives of people diagnosed with ASD in the Spanish population. BAPQ items were grouped into three subscales, which corresponded with the original BAP core factors (Hurley et al., 2007; Piven et al., 1997a, 1997b): aloofness, rigidity, and pragmatic language problems. The correspondence of the items and the subscales as well as the reversed items can be found in Hurley et al. (2007).

The Autism Quotient. The AQ (Baron-Cohen et al., 2001) was a 50-item self-report scale designed to identify adults diagnosed with high-functioning autism and standard intelligence. AQ items were grouped into five sub-scales: Social Skill, Attention Switching and Attention to detail, Communication, and Imagination. However, the test only provides a total score; the correspondence of the items and the subscales as well as the reversed items can be found in Baron-Cohen et al. (2001). The items of the AQ were adapted to the Spanish language following the guidelines proposed by Muñiz et al. (2013) and the International Test Commission Guidelines on Adapting a Test (http://www.intestcom.org) with the assistance of an official translator.

The Social Responsiveness Scale-2 (Adult Self-report Version). The SRS-2 (Constantino & Gruber, 2012) was a 65-item ordinally quantitative test for examining the severity of autistic traits in adults. SRS-2 items were grouped into five factors: Social Awareness, Social Cognition, Social Communication, Social Motivation, and Restricted Interests and Repetitive Behaviours; the correspondence of the items and the subscales, as well as the reversed items, can be found in Constantino and Gruber (2012). It also included two subscales compatible with the DSM-5: Social Communication and Interaction, and Restricted Interests and Repetitive Behaviours. As described previously, the items of the SRS-2 were adapted to the Spanish language following the guidelines proposed by Muñiz et al. (2013) and the International Test Commission Guidelines on Adapting a Test (http://www.intestcom.org) with the assistance of an official translator. At the time of conducting this study, there was no Spanish adaptation of the test available; the official Spanish SRS-2 adult self-form was published one year later (Constantino, 2017).

Item Selection and Assignment According to their Content and Itemmetric Properties

Assignment of Items to ASD/BAP Subdomains. Experts received two text documents by email, the first of which listed the seven subdomains of the updated BAP operationalization and their definitions (see Appendix 2, Table 1), and a second that included the items of the three questionnaires in this order: BAPQ, SRS-2, and AQ.

Itemmetric Analysis. Later, the experts also received an excel document containing the results from their previous assignment of the items. On this occasion, the selected items of the three questionnaires were randomized and included as a common pool of 121 items (39 items from the AQ, 36 items from the BAPQ, and 46 items from the SRS-2) without any reference to the original factor or questionnaire to which they belonged. The experts had to rate from 1 to 4 (for example, 1 = *not* *clear*, 2 = *somewhat clear*, 3 = *quite clear*, and 4 = *very clear*) each item according to the following properties: clarity (item was accurate and excluded double negations, adverbs incongruent with the rating scale, and multidimensionality), comprehensibility (readers could understand the item at the outset), concreteness (each item referred to only one idea), degree of self-reference (responses to the item could be expressed based on a person's perception of him/herself), and evaluation of the items (responses to the item could be influenced by social desirability).

Expert Judgment

The documentation for Expert Judgment included (i) a cover letter with information on the research group and the scope of the study; (ii) the updated BAP operationalization around two core domains and seven subdomains aligned with the updated ASD definition in the DSM-5 (APA, 2013; see Appendix 2, Table 1); (iii) a brief explanation and examples of BAP domains and subdomains. It also contained the instructions and variables targeted in this study (all files included in Expert judgment 2 can be consulted in Appendix 3. Document 1 and Appendix 4. Document 2). Following authors' recommendations (Almanasreh et al., 2019; Lynn, 1986) quantitative findings of content validity were collected for assessing the relevance of each item for the subdomain and representativeness of the BAP construct. Both of these aspects were rated by the experts using a Likert-type ordinal scale with four possible responses (for relevance: 1 = *not relevant*, 2 = *somewhat relevant*, 3 = *quite relevant*, and 4 = *very relevant*; and for representativeness: 1 = *very poorly represented*, 2 = *poorly represented*, 3 = *well represented*, and 4 = *very well represented*). Finally, we also took into account other variables beyond the scope of this study (experts were asked about the adequacy of the items to the objectives of a new test and whether the items were susceptible to differential functioning between targeted populations).

Procedure

Parallel Analysis

Two booklets containing the items from the AQ, the BAPQ, and the SRS-2, together with two other questionnaires beyond the scope of this study[9], were administered to the sample of students. Although the participants were given as much time as they needed to complete the questionnaires, the testing phase lasted approximately 50 min (for further information, see Godoy-Giménez et al., 2018).

Item Selection and Assignment According to their Content and Itemmetric Properties

First, a group meeting was held where the collaborators were given prior instructions about the entire procedure, the theoretical foundations of each questionnaire, and the updated BAP operationalization[10]. Then, each member independently allocated the items to the seven subdomains of that updated BAP operationalization. They were also asked to make a note of items that did not fit into any subdomain. Later, the five collaborators shared their ideas and discussed any discrepancies about item assignations. Finally, they independently assessed the itemmetric properties of all the items of the three questionnaires.

Expert Judgment

The documentation was sent to the experts by ordinary mail. They were informed that they had two weeks to complete the whole task and return it to us by a pre-addressed postage-paid envelope. The experts were assured that they could rely on our assistance during the assessment process, although none of them required it.

Data Analysis

Parallel Analysis

[9] In the present study, we used the same sample as that in Godoy-Giménez et al. (2018), where the Spanish version of the BAPQ was adapted and validity evidence was provided based on BAPQ relationships with measures of other variables: Pragmatic Awareness Questionnaire (PAQ; Rodríguez-Muñoz, 2012) and The NEO Five-Factor Inventory (NEO-FFI, Spanish Version; Costa & McCrae, 1999).

[10] Deficits in social-emotional reciprocity; Deficits in non-verbal communicative behaviours used for social interaction; Deficits in developing, maintaining, and understanding relationships; Stereotyped or repetitive motor movements, use of objects, or speech; Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behaviour; Highly restricted, fixated interests; and Hyper- or hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment (APA, 2013).

We conducted a parallel analysis of principal components (Horn 1965) on the total raw scores of each of the AQ, BAPQ, and SRS-2 subscales. The analyses were conducted using the Psych package (Psych 1. 9.12.31, 2020) in R software version 3.6.3 (R Core Team, 2020).

Item Selection and Assignment According to their Content and Itemmetric Properties

Following previously established guidelines (Angleitner et al., 1986; Osterlind, 1998) the items were selected according to their itemmetric properties; thus, only items that showed expert agreement (that is, with averages ≥ 3.8) on all the variables assessed were incorporated into the test. Since these questionnaires have been used to discriminate among people with different levels of BAP severity (e.g., Baron-Cohen et al., 2001; Constantino & Todd, 2003; Hoekstra et al., 2007; Sasson, et al., 2013b; Shi et al., 2015), or for predicting BAP-related variables (e.g., Faso et al., 2016; Hus et al., 2013; Sasson, et al., 2013c; Stewart et al., 2009; Takei et al., 2014), we considered that many of their items could serve to assess the BAP traits according to its updated definition.

Expert Judgment

Following Waltz and Bausell (1983), we took the proportion of items that received a rating of 3 or 4 by the experts as the content validity index (CVI; Hambleton et al., 1978). First, we collapsed four ordinal rating-scales into two dichotomous categories and assigned them labels (0 = *content invalid*, 1 = *content valid*). An expert agreement of 80% (indicating valid content) was taken as an index of item inclusion. Final decisions on items (i.e., conserving or deleting) were based on both data analysis and comments from the experts.

Equally, we examined (i) whether the construct of BAP was well represented by the domains and subdomains included in BAP operationalization and (ii) whether the items included in the questionnaire were sufficient to represent the construct of BAP. Furthermore, we asked the experts to suggest BAP characteristic behaviours or items that they considered relevant for improving the representation of the construct or any of its domains and subdomains.

Finally, once we had selected the items in terms of relevance to each subdomain, we verified whether the remaining items covered the entire theoretical content within the structure of BAP.

Specifically, (i) the relevance of the content of each item of the three questionnaires was studied according to the subdomain included in the BAP operationalization that showed the greatest content-based relationship, and (ii) the representativeness of a BAP test depended on the degree to which its items were proportionally distributed or weighted across the two core domains and seven subdomains (as conceptualized in alignment with the ASD definition in the APA, 2013) and whether items may cover the entire BAP construct. Considering the DSM-5 specifications of ASD deficits (APA, 2013) and the absence of any work which has highlighted that some autism deficits are more pivotal than others, we considered that the seven subdomains should be adequately represented in the final pool of items. The reason why we performed two subsequent expert judgments was due to the huge number of items. That way, a first group allocated the items and screened them to select only those with adequate psychometric properties while the second broader group confirmed that item assignment was correct and that the BAP construct was correctly represented by the items.

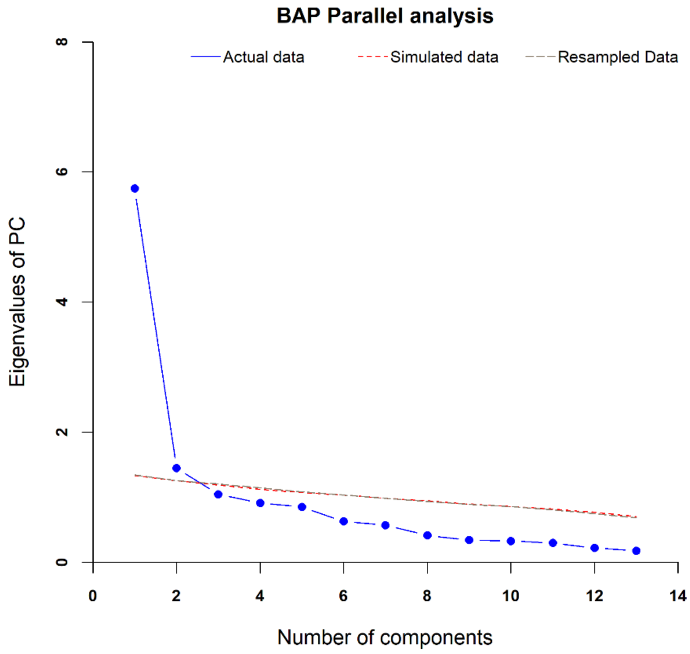
Results

Parallel Analysis

The results of the Parallel analysis of principal components (see Figure 1) lead us to propose that a solution of two-components underlies the covariation of the 13 BAP subscales. Furthermore, most of the variance is explained by one of these two components. The first component explained 44.23% of the total variance while the second component explained 11.15%, together they explained 55.39% of the total variance.

Figure 1

Results of Parallel Analysis Principal Components



Note. Actual data = original data; Simulated data = random data with the same N variables and sample size; Resampled Data = repeated sample from the original data.

Item Selection and Assignment According to their Content and Itemmetric Properties

Assignment of Items According to their Content

Assignment of the items to each of the seven ASD subdomains is displayed in Figure 2. Likewise, specific items are presented in Appendix 2, Table 2. A total of 83 items were found to fit the SCI BAP domain while 38 fitted the RRB BAP one.

Closer inspection of the data revealed that there was also an imbalance in the distribution of the items among the subdomains. In this regard, two subdomains of RRB BAP— Stereotyped or repetitive motor movements and Hyper- or hypo-reactivity to sensory input or unusual interests, accounted for only a few items, primarily from the SRS-2. With regard to discarded items, all BAPQ items were found to fit a subdomain while 11 AQ items and 19 SRS-2 items were eliminated.

Itemmetric Analysis

Results from the itemmetric analysis are also shown in Figure 2, Appendix 2, Tables 2, and 3. Forty-six items were eliminated by the experts (this time the discarded items belonged equally to the three questionnaires), due to a lack of clarity, comprehensibility, or concreteness, degree of self-reference, and/or the evaluation of the items. However, after reviewing the results, the experts decided to keep two items of the SRS-2 (thus 44 items were finally discarded) due to the lack of representation in Highly restricted, fixated interests, since the experts considered their contents to be key for the subdomain. Finally, RRB BAP was still the domain with the least number of items (62 SCI BAP/21 RRB BAP).

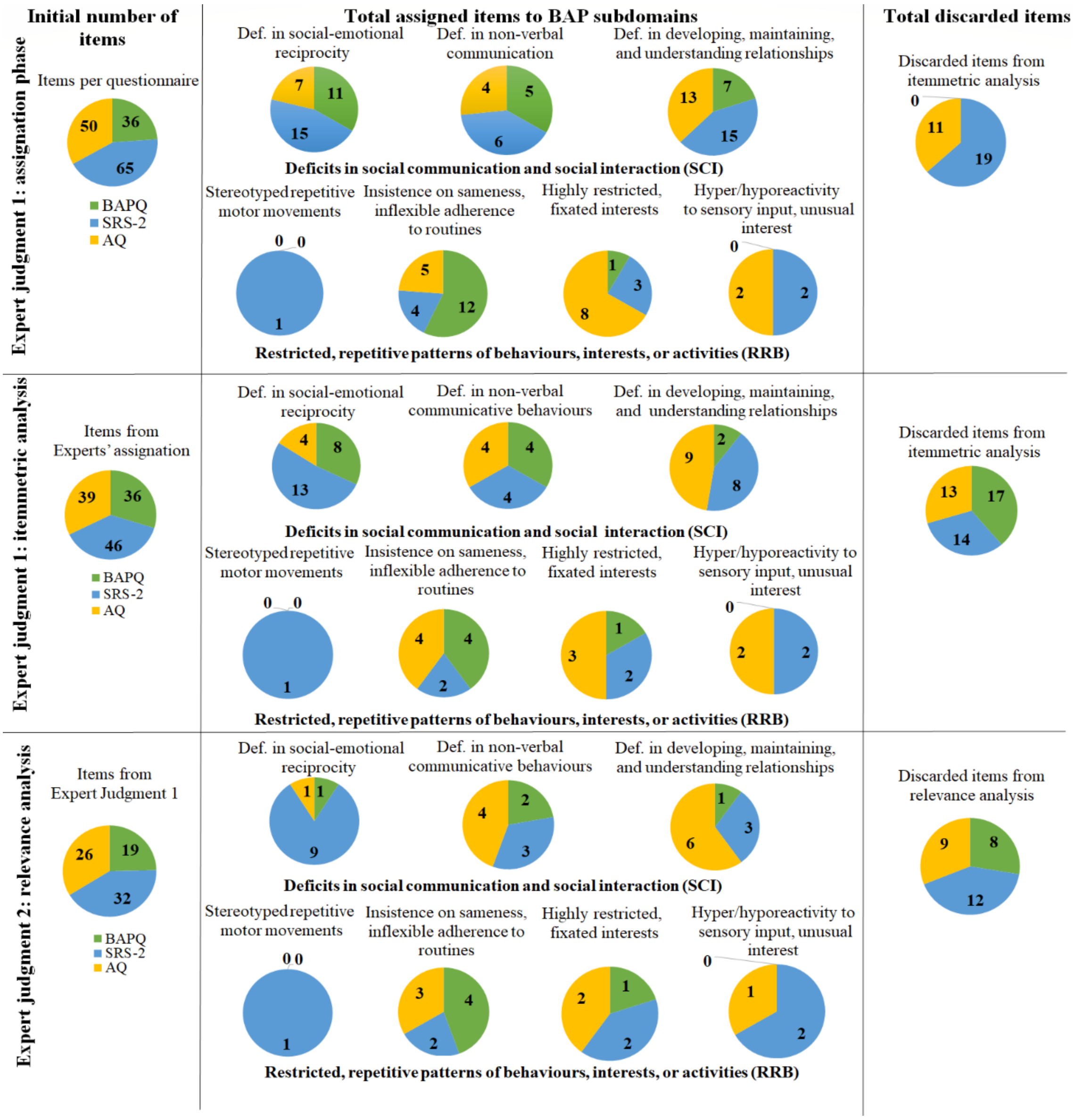
Expert Judgment

The Relevance of the Items for the Subdomains

Results from Expert judgment 2 (see Figure 2; see Appendix 2, Table 2) revealed a remaining pool of 48 items according to their relevance to the subdomain, 30 for SCI BAP, and 18 for RRB BAP. Stereotyped or repetitive motor movements and Hyper- or hypo-reactivity to sensory input or unusual interest subdomains in RRB BAP still accounted for few items but were maintained while the remaining subdomains of both domains now had fewer items. Twenty-nine items were discarded by the experts and these were distributed equally across the three questionnaires.

Figure 2

Results from both Expert Judgments



Representativeness of the BAP Construct

Fifteen percent of the experts rated the representation of the BAP construct as very well represented and 85% considered it well represented. Similarly, 95% of the experts considered that the items of the questionnaire represented the construct well; whilst one response was missing.

Discussion

The purpose of this study was to connect the content of three current BAP measurement tools with an updated operationalization of the BAP harmonized with the ASD one (APA, 2013). The underlying idea was to detect possible gaps in the contents of BAP measurement instruments and to take a preliminary step towards bridging the gap between the operationalization and measurement of this phenotype.

First, we conducted a parallel analysis (Horn, 1965) to assess the factor structure of the BAP across the AQ, the BAPQ, and the SRS-2 subscales (Hurley et al., 2007; Piven, et al., 1997a, 1997b). Unlike the findings reported by Wainer et al. (2011), our results pointed to two main components instead of three. The fact that the three most applied BAP tests can cluster the variance of their 13 scales into two components provides an empirical basis for updating the BAP operationalization to a construct that uses two principal domains aligned with the current definition of ASD (APA, 2013). These results are consistent with those presented previously providing empirical evidence for the correspondence between BAP and ASD dimensions. For example, Sasson and et al., (2013c) observed that socio-communicative aspects of the BAP (measured by the Aloof and Pragmatic Language subscales of the BAPQ), unlike non-social features (measured by the BAPQ Rigid subscale) were connected to social skill and social cognition whilst also predicting poorer performance in social areas. In another study, Frazier et al. (2014) examined the SRS-2 factor structure and considered a simpler two-factor solution that showed the correspondence between SRS-2 and current ASD dimensions.

Second, we explored whether the available self-reported tests were adequate for evaluating BAP according to its updated definition. The study of the phenotype requires an adequate measurement tool that cannot disregard the actual structure and specifications of the construct. Thus, both, theoretically and empirically-based claims regarding the two-dimensional structure of the BAP (e.g., Morrison et al., 2018; Sasson, et al., 2013c) formed the basis of the main goal of this study: to reallocate its items (i.e., relevant BAP behaviours or preferences) according to the two main subdomains of ASD in the DSM-V (APA, 2013).

With respect to this allocation phase, two main observations are worth noting. All the BAPQ items were distributed across different subdomains of the proposed BAP operationalization but they did not sufficiently cover all the subdomains. That is, whilst BAPQ items might accurately evaluate subclinical autistic-like traits according to the DSM-5, the inferences based only on its items would be limited and not representative of the actual BAP construct. The SRS2, however, with its higher number of items and contents, taps into all BAP subdomains and gathers more types of BAP-related behaviours and preferences than the BAPQ and the AQ. However, at the same time, some SRS-2 items were considered to not fit in any subdomain. Consequently, applying only the SRS-2 when measuring autism traits could cover all autistic-like features, together with some other non-autistic related deficits, which could blur the interpretation of its scores. On balance, we conclude that, when used alone, all of the proposed tests are insufficient for assessing the updated BAP operationalization. Consequently, the next step was to test whether a conjoint use of the items of the three tests would more adequately represent the BAP construct.

Third, before studying the relevance and representativeness of all these items for measuring BAP (Expert judgment 2), their wordings were scrutinized to ensure that any problem regarding their itemmetric properties would not affect the work of the experts. It is worth highlighting the importance of eliminating poorly drafted items that could yield biased responses since these could affect the way respondents interpret and answer them (see Angleitner et al., 1986; Grant & Davis, 1997), along with the way the construct is finally assessed (Haynes et al., 1995). Furthermore, some populations, particularly BAP people, have problems when they are required to put themselves in the position of others. Consequently, they may struggle to answer items regarding others' view of the person (non-self-referred items). As we hypothesized, some formal problems emerged during the itemmetric judgment, and half of the items were discarded in this analysis. The item property that appeared to be the most problematic was concreteness. The experts suggested that some items were too generic and/or refer to more than one idea (e.g., SRS-2 item 30 refers to more than one idea) whilst others were unclear (e.g., AQ Item 15 uses an adverb that interferes with the rating scale) or did not involve self-reference.

Fourth, we examined the relevance and representativeness of the items for covering the whole BAP construct (Haynes et al., 1995; Lynn, 1986; Nunnally & Bernstein, 1994; Suen & Ary, 1989). By doing this, the initial set of items was reduced by less than a third. In addition, both expert judgments indicated an unbalance in the distribution of items throughout the domains and subdomains where the number of items in SCI BAP was more than twice those in RRB BAP. This unbalance could be due to discrepancies in the original operationalizations of the targeted construct of each test. The BAPQ was developed based on a three-dimensional BAP structure (Hurley et al., 2007), while the AQ also included aspects of cognitive abnormality (Baron-Cohen et al., 2001). Similarly, the SRS originally aimed to assess autistic social impairment, particularly reciprocal social behaviours (Constantino & Todd, 2005) although the most recent version of this questionnaire, the SRS-2, includes a new subscale for the assessment of restricted and repetitive behaviours (Constantino & Gruber, 2012).

This asymmetric representation of domains could be challenging, given that many authors have claimed that both are important for diagnosing autism since the time that Kanner (1943) provided the first accurate definition of the disorder. Accordingly, and as now detailed in the DSM-5, the three SCI subdomains and at least two RRB subdomains must hinder the person's everyday life to be regarded as a conclusive ASD diagnostic (APA, 2013). The same can be applied to BAP assessment such as the non-clinical expressions within the autism spectrum. Furthermore, the experts (Expert judgment 2) in our study considered that the BAP construct was well represented by the two domains and seven subdomains included in the definition we provided based on the autism spectrum definition (DSM-5; APA, 2013). These results support our claim for the need to evaluate both SCI and RRB domains equally when studying BAP.

The experts also pointed out that the selected items were representative of the construct. Although the content of the items could appear to be sufficiently representative of the BAP construct, the final allocation presents clear gaps in the BAP content. Thus, we should not disregard the underrepresentation of two RRB subdomains and their theoretical and psychometric implications. Since autism shares some indicators with other disorders, the worst scenario that could arise from neglecting certain key autistic behaviours (e.g., Hyper/hypo-reactivity to sensory input or unusual interest in sensory aspects) in preliminary test construction phases could lead to variations in final test scores that only reflect differences in traits that are also shared with other disorders.

In this regard, some studies have reported that the RRB domain can be divided into two clusters of indicators: (a) repetitive motor and sensory behaviours (repetitive hand movements) and (b) insistence on sameness (narrow interests, rigid routines, and rituals; Cuccaro et al., 2003; Honey et al., 2012; Richler et al., 2007). Assessing only one of the two RRB subtypes could lead to ASD variations being confounded with other disorders such as social communication disorder (characterised by persistent deficits in the social use of verbal and non-verbal communication in the absence of restricted and repetitive interests and behaviours; APA, 2013) or obsessive-compulsive disorder (the assessed person could not meet the criteria for the second subdomain). This would have a direct effect on BAP identification and future research studies, particularly those concerned with neurobiological and genomic aspects (Ruscio & Ruscio, 2002, 2004). For instance, there is evidence that the subdomains of RRB are underpinned by different neural pathways (Langen et al., 2011). Thus, a lack of relevant items with which to assess stereotyped or repetitive motor movements, use of objects, or speech, could negatively affect the study of the different neural pathways of autism, either in people diagnosed with autism or BAP family members.

In conclusion, researchers have begun to highlight the need to update the definition of BAP so that it is aligned with the current definition of ASD (Morrison et al., 2018; Sasson, et al., 2013c). This study represents a first step towards achieving this goal by providing empirical evidence in support of the need for a new test for evaluating the BAP that runs parallel to the ASD structure, containing its most relevant content but also including additional indicators that measure milder forms of ASD.

Supplementary Information. The online version contains supplementary material available at https://doi.org/10.1007/s10803-021-05158-7

Author Contribution. All the authors took part in the Item Selection and Assignment According to their Content and Itemmetric Proper-ties. MGG and AGR were responsible for data collection. PSJ, AFE, and FC were responsible for the design of the study. PSJ and MGG analysed the data and wrote the manuscript. All authors contributed to data interpretation and assisted with writing the manuscript. All the authors gave their consent for publishing this study.

Funding. This work was supported by grants from the Spanish Ministry of Science and Innovation (Agencia Estatal de Investigación, AEI) to AFE (PID2019-110066GB-I00/AEI/10.13039/501100011033); and the Spanish Ministry of Education, Culture, and Sport to MGG (FPU2015-015).

Declarations

Conflict of interest. The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethical Approval. This study received ethics approval from the local Human Research Ethics Committees of the University of Almeria (UALBIO2017/006) involved in the recruitment process which was conducted under the approved guidelines and the Declaration of Helsinki.

Informed Consent. Written informed consent was obtained from all undergraduates (parallel analysis) before conducting the study. They received two-course credits for their collaboration. We reported how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

References

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, *15*(2), 214-221. https://doi.org/10.1016/j.sapharm.2018.03.066

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, DC: Joint Committee on Standards for Educational and Psychological Testing.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. Washington, DC: American Psychiatric Pub.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders revised (DSM-IV-TR).* Washington, DC: American Psychiatric Pub.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (DSM-IV).* Washington, DC: American Psychiatric Pub.

Angleitner A., John O.P., Löhr FJ. (1986) It's *What* You Ask and *How* You Ask It: An Itemmetric Analysis of Personality Questionnaires. In: Angleitner A., Wiggins J.S. (eds) *Personality Assessment via Questionnaires.* Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-70751-3_5

Armstrong, T. S., Cohen, M. Z., Eriksen, L., & Cleeland, C. (2005). Content Validity of Self-Report Measurement Instruments: An Illustration From the Development of the Brain Tumor Module of the M.D. Anderson Symptom Inventory. *Oncology Nursing Forum, 32*(3), 669–676. https://doi.org/10.1188/05.onf.669-676

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists, and mathematicians. *Journal of Autism and Developmental Disorders, 31,* 5–17. https://doi.org/10.1023/A:1005653411471

Beck, C. T., & Gable, R. K. (2001). Ensuring content validity: An illustration of the process. *Journal of Nursing Measurement*, *9*(2), 201-215. https://doi.org/10.1891/1061-3749.9.2.201

Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., ... & Rutter, M. (1994). A case-control family history study of autism. *Journal of child Psychology and Psychiatry, 35*(5), 877-900. https://doi.org/10.1111/j.1469-7610.1994.tb02300.x

Bruni, T. P. (2014). Test Review: Social Responsiveness Scale–Second Edition (SRS-2) ConstantinoJ. N.GruberC. P. (2012). Social Responsiveness Scale–Second Edition (SRS-2). Torrance, CA: Western Psychological Services. *Journal of Psychoeducational Assessment, 32*(4), *365–369.* https://doi.org/10.1177/0734282913517525

Constantino, J. N., & Gruber, C.P. (2005). *The Social Responsiveness Scale™.* Torrance, CA: Western Psychological Services.

Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale: SRS-2.* Torrance, CA: Western Psychological Services.

Constantino, J. N., Lajonchere, C., Lutz, M., Gray, T., Abbacchi, A., McKenna, K., . . . Todd, R. D. (2006). Autistic social impairment in the siblings of children with pervasive developmental disorders. *The American Journal of Psychiatry, 163*, 294–296. https://doi.org/10.1176/appi.ajp.163.2.294

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Archives of general psychiatry, 60*, 524-530. http://dx.doi.org/10.1001/archpsyc.60.5.524

Cuccaro, M.L., Shao, Y., Grubber, J. et al. (2003). Factor Analysis of Restricted and Repetitive Behaviors in Autism Using the Autism Diagnostic Interview-R. *Child Psychiatry Hum Dev, 34*, 3–17. https://doi.org/10.1023/A:1025321707947

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied nursing research*, *5*(4), 194-197. https://doi.org/10.1016/S0897-1897(05)80008-4

Delgado-Rico, E., Carretero-Dios, H., & Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology España*, *12*(3), 449-460. https://doi.org/10.5167/uzh-64551

Ebel, R. L., & Frisbie, D. A. (1972). *Essentials of educational measurement* (pp. 492-494). Englewood Cliffs, NJ: Prentice-Hall.

English, M. C., Gignac, G. E., Visser, T. A., Whitehouse, A. J., & Maybery, M. T. (2019). A comprehensive psychometric analysis of autism-spectrum quotient factor models using two large samples: Model recommendations and the influence of divergent traits on total-scale scores. *Autism Research.* https://doi.org/10.1002/aur.2198

Faso, D. J., Corretti, C. A., Ackerman, R. A., & Sasson, N. J. (2016). The broad autism phenotype predicts relationship outcomes in newly formed college roommates. *Autism, 20*(4), 412-424. https://doi.org/10.1177/1362361315585733

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2014). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism, 18*, 31–44. https://doi.org/10.1177/1362361313500382

Gable, R. K., & Wolf, M. B. (2012). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings* (Vol. 36). Springer Science & Business Media. https://doi.org/10.1007/978-94-011-1400-4

Godoy-Giménez, M., González-Rodríguez, A., Cañadas, F., Estévez, A. F., & Sayans-Jiménez, P. (2018). Psychometric Properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements. Journal of autism and developmental disorders, 1-14. https://doi.org/10.1007/s10803-017-3438-0

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in nursing & health, 20*(3), 269-274. https://doi.org/10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.0.CO;2-G

Hambleton, R. K., Swaminathan, H., Algina, J., & Coulson, D. B. (1978). Criterion-referenced testing and measurement: A review of technical issues and developments. *Review of Educational Research*, *48*, 1-47. https://doi.org/10.3102/00346543048001001

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods.

Psychological Assessment, 7(3), 238–247. https://doi.org/10.1037/1040-3590.7.3.238

Hoekstra, R. A., Bartels, M., Verweij, C. J., & Boomsma, D. I. (2007). Heritability of autistic traits in the general population. *Archives of Pediatrics & Adolescent medicine, 161*(4), 372-377. http://dx.doi.org/10.1007/s10803-008-0538-x

Honey, E., Rodgers, J., & McConachie, H. (2012). Measurement of restricted and repetitive behaviour in children with autism spectrum disorder: Selecting a questionnaire or interview. *Research in Autism Spectrum Disorders, 6*(2), 757-776. https://doi.org/10.1016/j.rasd.2011.10.011

Hurley, R. S., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of autism and developmental disorders, 37,* 1679-1690. http://dx.doi.org/10.1007/s10803-006-0299-3

Hurst, R. M., Mitchell, J. T., Kimbrel, N. A., Kwapil, T. K., & Nelson-Gray, R. O. (2007). Examination of the reliability and factor structure of the Autism Spectrum Quotient (AQ) in a non-clinical sample. *Personality and Individual Differences, 43*(7), 1938-1949. http://dx.doi.org/10.1016/j.paid.2007.06.012

Hus, V., Bishop, S., Gotham, K., Huerta, M., y Lord, C. (2013). Factors influencing scores on the social responsiveness scale. *Journal of Child Psychology and Psychiatry, 54*(2), 216–224. https://doi.org/10.1111/j.1469-7610.2012.02589.x

Ingersoll, B., & Wainer, A. (2014). The broader autism phenotype. *Handbook of Autism and Pervasive Developmental Disorders, Fourth Edition*.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250.

Langen, M., Durston, S., Kas, M. J., van Engeland, H., & Staal, W. G. (2011). The neurobiology of repetitive behavior:… and men. *Neuroscience & Biobehavioral Reviews, 35*(3), 356-365. https://doi.org/10.1016/j.neubiorev.2010.02.005

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3,* 635-694. https://doi.org/10.2466/pr0.1957.3.3.635

Losh, M., Childress, D., Lam, K., & Piven, J. (2008). Defining key features of the broad autism phenotype: A comparison across parents of multiple-and single-incidence autism families. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 147,* 424–433. https://doi.org/10.1002/ajmg.b.30612

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017

Morrison, K. E., Chambers, L. K., Faso, D. J., & Sasson, N. J. (2018). The content and function of interests in the broad autism phenotype. *Research in Autism Spectrum Disorders*, *49,* 25-33. http://dx.doi.org/10.1016/j.rasd.2018.02.002

Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Guidelines for test translation and adaptation: Second edition [Directrices para la traducción y adaptación de los tests: Segunda edición]. *Psicothema, 25*, 151–157.

Murphy, M., Bolton, P. F., Pickles, A., Fombonne, E., Piven, J., & Rutter, M. (2000). Personality traits of the relatives of autistic probands. *Psychological Medicine, 30*(6), 1411-1424. http://dx.doi.org/10.1017/S0033291799002949

Narayan, S., Moyes, B., & Wolff, S. (1990). Family characteristics of autistic children: A further report. *Journal of Autism and Developmental Disorders, 20*(4), 523-535. https://doi.org/10.1007/BF02216057

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGrawHill.

Osterlind, S. J. (1998). *What Is constructing test items?* (pp. 1-16). Springer Netherlands.

Piven, J., & Palmer, P. (1999). Psychiatric disorder and the broad autism phenotype: evidence from a family study of multiple-incidence autism families. American *Journal of Psychiatry, 156*(4), 557-563. https://doi.org/10.1176/ajp.156.4.557

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997b). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry, 154*(2), 185–190. http://dx.doi.org/10.1176/ajp.154.2.185

Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997a). Personality and language characteristics in parents from multiple-incidence autism families. American Journal of Medical Genetics, 74(4), 398-411. https://doi.org/10.1002/(SICI)1096-8628(19970725)74:4<398::AID-AJMG11>3.0.CO;2-D

Piven, J., Wzorek, M., Landa, R., Lainhart, J., Bolton, P., Chase, G. A., & Folstein, S. (1994). Personality characteristics of the parents of autistic individuals. *Psychological Medicine, 24*(3), 783-795. https://doi.org/10.107/S0033291700027938

Richler, J., Bishop, S.L., Kleinke, J.R. et al. (2007). Restricted and Repetitive Behaviors in Young Children with Autism Spectrum Disorders. *J Autism Dev Disord, 37*, 73–85. https://doi.org/10.1007/s10803-006-0332-6

Ruscio, J., & Ruscio, A. M. (2002). A structure-based approach to psychological assessment: Matching measurement models to latent structure. *Assessment*, *9*, 4-16. https://doi.org/10.1177/1073191102091002

Ruscio, J., & Ruscio, A. M. (2004). Clarifying boundary issues in psychopathology: the role of taxometrics in a comprehensive program of structural research. *Journal of abnormal psychology*, *113*, 24. https://doi.org/10.1037/0021-843X.113.1.24

Russell-Smith, S. N., Maybery, M. T., & Bayliss, D. M. (2011). Relationships between autistic-like and schizotypy traits: An analysis using the Autism Spectrum Quotient and Oxford-Liverpool Inventory of Feelings and Experiences. *Personality and Individual Differences, 51*(2), 128–132. https://doi.org/10.1016/j.paid.2011.03.027

Sasson, N. J., Lam, K. S. L., Childress, D., Parlier, M., Daniels, J. L., & Piven, J. (2013a). The Broad Autism Phenotype Questionnaire: Prevalence and Diagnostic Classification. *Autism Research, 6,* 134–143. http://dx.doi.org/10.1002/aur.1272

Sasson, N. J., Lam, K. S., Parlier, M., Daniels, J. L., & Piven, J. (2013b). Autism and the broad autism phenotype: familial patterns and intergenerational transmission. *Journal of Neurodevelopmental Disorders, 5,* 11. http://dx.doi.org/10.1186/1866-1955-5-11

Sasson, N. J., Nowlin, R. B., & Pinkham, A. E. (2013c). Social cognition, social skill, and the broad autism phenotype. *Autism, 17,* 655-667. http://dx.doi.org/10.1177/1362361312455704

Seidman, I., Yirmiya, N., Milshtein, S., Ebstein, R. P., & Levi, S. (2012). The Broad Autism Phenotype Questionnaire: mothers versus fathers of children with an autism spectrum disorder. *Journal of autism and developmental disorders, 42*(5), 837-846. http://dx.doi.org/10.1007/s10803-011-1315-9

Shi, L.-J., Ou, J.-J., Gong, J.-B., Wang, S.-H., Zhou, Y.-Y., Zhu, F.-R., ... Luo, X.-R. (2015). Broad autism phenotype features of Chinese parents with autistic children and their associations with severity of social impairment in probands. *BMC Psychiatry, 15.* http://dx.doi.org/10.1186/s12888-015-0568-9

Stewart, M. E., & Austin, E. J. (2009). The structure of the Autism-Spectrum Quotient (AQ): Evidence from a student sample in Scotland. *Personality and Individual Differences, 47*, 224-228. http://dx.doi.org/10.1016/j.paid.2009.03.004

Stewart, M. E., Watson, J., Allcock, A. J., & Yaqoob, T. (2009). Autistic traits predict performance on the block design. *Autism, 13*(2), 133-142. https://doi.org/10.1177/1362361308098515

Suen, H., & Ary, D. (1989). *Analyzing quantitative observation data*. Hillsdale, NJ: Lawrence Erlbaum.

Takei, R., Matsuo, J., Takahashi, H., Uchiyama, T., Kunugi, H., & Kamio, Y. (2014). Verification of the utility of the social responsiveness scale for adults in non-clinical and clinical adult populations in Japan. *BMC psychiatry, 14,* 302. https://doi.org/10.1186/s12888-014-0302-z

Tilden, V. P., Nelson, C. A., & May, B. A. (1990). Use of qualitative methods to enhance content validity. *Nursing Research, 39*(3), 172–175. https://doi.org/10.1097/00006199-199005000-00015

Wainer, A. L., Ingersoll, B. R., & Hopwood, C. J. (2011). The Structure and Nature of the Broader Autism Phenotype in a Non-Clinical Sample. *Journal of Psychopathology and Behavioral Assessment, 33*, 459–469. http://dx.doi.org/10.1007/s10862-011-9259-0

Waltz, C., & Bausell, R. B. (1983). *Nursing Research: design, statistics, and computer analysis.* Philadelphia: FA Davis Company.

Wheelwright, S., Auyeung, B., Allison, C., & Baron-Cohen, S. (2010). Defining the broader, medium, and narrow autism phenotype among parents using the Autism Spectrum Quotient (AQ). *Molecular Autism, 1,* 1–10. https://doi.org/10.1186/2040-2392-1-10

Whitehouse, A. J. O., Coon, H., Miller, J., Salisbury, B., & Bishop, D. V. M. (2010). Narrowing the broader autism phenotype: A study using the Communication Checklist—Adult Version (CC-A). *Autism, 14*(6), 559–574. https://doi.org/10.1177/1362361310382107

CHAPTER SIX. STUDY 4:

THE BROAD AUTISM PHENOTYPE-INTERNATIONAL TEST (BAP-IT): A TWO-DOMAIN BASED TEST FOR THE ASSESSMENT OF THE BROAD AUTISM PHENOTYPE

Manuscript being prepared as:

Godoy-Giménez, M., González-Rodríguez, A., García-Pérez, A., Estévez, A. F., Cañadas, F., & Sayans-Jiménez, P. (2022). The Broad Autism Phenotype - International Test (BAP-IT): a two-domain based test for the assessment of the broad autism phenotype

The Broad Autism Phenotype - International Test (BAP-IT): a Two-Domain Based Test for the Assessment of the Broad Autism Phenotype

Marta Godoy-Giménez[1], Antonio González-Rodríguez[1], Ángel García-Pérez[1], Ángeles F. Estévez[1,2], Fernando Cañadas[1,2], and Pablo Sayans-Jiménez[1]

[1] Department of Psychology, University of Almería, Almería, Spain

[2]CERNEP Research Centre, University of Almería, Almería, Spain

Author Note

Marta Godoy-Giménez https://orcid.org/0000-0002-2298-0113

Antonio González-Rodríguez https://orcid.org/0000-0001-6432-829X

Angel García-Pérez https://orcid.org/0000-0001-6375-6823

Angeles F. Estévez http://orcid.org/0000-0003-0067-0840

Fernando Cañadas http://orcid.org/0000-0001-7689-0132

Pablo Sayans-Jiménez https://orcid.org/0000-0001-6681-5729

Correspondence concerning this article should be addressed to Pablo Sayans-Jiménez, Department of Psychology, University of Almería, 04120 Almería, Spain. Email: psj242@ual.es telephone number: +34 950 214797, and Fernando Cañadas, Department of Psychology, Cernep Research Centre, University of Almería, 04120 Almería, Spain. Email: jcanadas@ual.es

Abstract

The broad autism phenotype (BAP) comprehends the phenotypic expressions of the primary characteristics of autism spectrum disorder (ASD). Available tests do not agree with the operationalisation of BAP: their internal structures have been shown to be unstable, and none of them use BAP operationalisation revolving around two domains coherent with ASD. This work presents the development of an entirely new measurement tool, the Broad Autism Phenotype—International Test (BAP-IT), for assessing the two-domain operationalisation of BAP in Spain and the United Kingdom. Careful protocols for test and item construction and adaptation were followed in a broad and international pilot study. Common factor approaches were employed to find a simple and invariant structure across countries. We provide validity evidence based on the internal structure of the test, dimension consistency, and extensive evidence based on the connections of the BAP-IT dimensions with other related variables. Finally, the BAP-IT comprehends 20 indicators distributed in two dimensions, SOCIAL-BAP and RIRE-BAP; these 20 items are representative of the seven autistic subdomains. The BAP-IT counts on fully metric invariance between the Spanish groups and partial metric invariance across countries and good omega reliability estimations. It establishes expected and adequate connections that link SOCIAL-BAP and RIRE-BAP to the main domains of BAP and ASD. We conclude that the provided evidence shows adequate functioning of the BAP-IT in comparison to other available BAP tests.

Keywords: Broad Autism Phenotype; Autism Spectrum Disorder; Validity Evidence; Test Construction; Test Invariance

The Broad Autism Phenotype - International Test (BAP-IT): a Two-Domain Based Test for the Assessment of the Broad Autism Phenotype

The broad autism phenotype (BAP) refers to the phenotypic expression of elevated but nonclinical levels of autism spectrum disorder (ASD), which expands autism beyond the diagnostic threshold, passing by the relatives of individuals diagnosed with ASD towards the general population (Constantino & Todd, 2003; Hurst et al., 2007; Stewart & Austin, 2009). In light of the hypothesis of a spectrum of autism, researchers have claimed that the BAP shares the primary characteristics of ASD (e.g. Morrison et al., 2018; Nayar et al., 2021; Pruit et al., 2018; Wainer et al., 2011). Nevertheless, its operationalisation has not been revised and updated congruent with the current definition of ASD in the DSM–5 (American Psychiatric Association [APA], 2013) [11] and the forthcoming CID-11 (World Health Organization [WHO], 2019) until recently, when Godoy-Giménez et al. (2021a) addressed the issue. They updated the BAP operationalisation after thoroughly reviewing all of the theoretical and empirical evidence supporting the reconciliation of the BAP and ASD constructs (e.g. Morrison et al., 2018; Nayar et al., 2021; Wainer et al., 2011). In the present manuscript, the updated two-dimensional operationalisation of the BAP (Godoy-Giménez et al., 2021a) has been undertaken.

The reunion of the BAP and ASD operationalisations might highly contribute to this field of research. Among the potential benefits, it would insert the BAP in the autism spectrum, thus increasing its understanding and, also, providing further information regarding how core autistic deficits are expressed differentially on each severity level of the continuum (Godoy-Giménez et al., 2021a). As the BAP expressions have been explored in twins, siblings and parents of diagnosed individuals – as well as in community samples – it could offer a greater range of variability in symptomatology, thus increasing the richness in the characterisation of the whole spectrum. These samples could also serve to track the role of genetics in ASD aetiology (Gaugler et al., 2014), because studying the presence of BAP in the

[11] According to the DSM-5, ASD is defined by the presence of impaired social communication and social relation (criterion A; APA, 2013) together with a pattern of restricted, repetitive behaviours and interests (criterion B; APA, 2013). This operationalisation configures the necessity of a test with two main domains of content and, predictably, two differentiated dimensions formed by the item's scores regarding each kind of content.

parents of a child diagnosed with ASD could uncover a genetic mechanism (Robinson et al., 2011).

With the reunion of BAP and ASD, authors may also use BAP samples to conduct research that is scarcely possible to carry out in clinical populations (Landry & Chouinard, 2016), as in the case of neuroimaging studies. Additionally, BAP-based samples offer the opportunity to study autistic symptoms in isolation, without the involvement of other comorbid disorders such as intellectual disability (APA, 2013). Finally, it also helps in examining the developmental path of symptoms. In this regard, studies like Stewart et al. (2018) have indicated that elder adults with BAP may feel higher levels of depression and anxiety and have lower executive functions, which would justify the need to explore the BAP in elder adults.

However, despite all of the advantages it entails, the reconciliation of the BAP and ASD necessarily drives restoration of any measurement process, which, indeed, represents a preliminary stage in almost every further experimental study. The most commonly applied measures in the BAP field are self-reports as the preferred method in personality psychology (Ingersoll & Wainer, 2014; Vazire, 2006). Although this has plenty of benefits (e.g. reducing time costs, inexpensive, rapid data collection and enhanced objectivity), researchers should be especially careful, as not all self-reports were originally designed to measure the specific BAP construct (e.g. the Autism Quotient [AQ] measures autistic-like behaviours in adults with normal intelligence; see Table 1). Thus, they do not share the BAP operationalisation nor the same internal structure, which could seriously affect the inferences made on the scores of these tests (e.g. the Social Responsiveness Scale—2 [SRS-II] or the Subthreshold Autism Trait Questionnaire [SATQ] comprehends five main factors, whereas the Broad Autism Phenotype Questionnaire [BAPQ], which was originally designed to assess the BAP, counts on three components that parallel DSM-IV autistic domains; APA, 1994; see Table 1).

Table 1

Revision of the Commonly Applied Tests for Assessing the BAP

Test	Target construct	Target population	Ground operationalization	Pros	Cons
Broader Phenotype Autism Symptom Scale (BPASS; Dawson et al., 2007)	BAP	Affected children, unaffected siblings, and parents	<u>Four domains</u>: ❊Social Motivation ❊Social Expressiveness ❊Conversational Skills ❊Flexibility/range of Interests	❊Provides a quantitative measure of autistic traits ❊Interview richness ❊Qualitative data ❊Structure related to the autistic dimensions in the DSM- IV (APA, 1994).	❊Long interview ❊The same interview either for children or adults as the interviewer must accommodate the patient. ❊Require trained clinicians and direct observation. ❊Not been commonly applied. ❊Does not share the two-domain structure of the ASD.
Broad Autism Questionnaire (BAPQ; Hurley et al, 2007)	BAP	Parents (originally), general adults	<u>Three subscales</u>: ❊Aloofness ❊Pragmatic Language Deficits ❊Rigidity	❊Structure corresponding to each of the autistic dimensions in the DSM-IV (APA, 1994). ❊Suggested instrument to identified BAP (Hurley et al., 2007; Wainer et al., 2011). ❊Good reliability and some favourable validity evidence (Hurley et al., 2007; Ingersoll et al., 2011; Sasson et al. 2013a).	❊Inadequate validity evidence based on internal structure: dimensionality problems in Pragmatic Language Deficits (Sasson et al., 2013a) and comparative fit index (CFI) indicators far from acceptable (e.g., Broderick et al., 2015). ❊Unreported factor loadings (e.g., Broderick et al., 2015; Ingersoll et al., 2011; Wainer et al., 2011).

Test	Target construct	Target population	Ground operationalization	Pros	Cons
				❇Self- and informant-report versions. ❇Applied in the general population. ❇Commonly and worldwide applied to measure BAP. ❇Many translated versions (e.g., *Chinese*, Shi et al., 2015; *Hindi*, Meera et al., 2015; *Israel,* Seidman et al., 2012). ❇Free access to the original BAPQ (Hurley et al., 2007) and some adapted versions.	❇Suggestions of items redistribution into a two-factor solution (Wainer et al., 2011). ❇Does not share the two-domain structure of the ASD.
Subthreshold Autism Trait Questionnaire (SATQ; Kanne et al., 2012)	BAP	General adults	<u>Five factors</u>: ❇Social Interaction & Enjoyment ❇Oddness ❇Reading Facial Expressions	❇Brief self-report questionnaire. ❇Grasps relevant expressions in broader populations. ❇Differentiate between ASD and common students.	❇Unacceptable fit for one and three-factor alternatives (Kanner et al., 2012). ❇Not very commonly applied. ❇Does not share the two-domain structure of the ASD.

Test	Target construct	Target population	Ground operationalization	Pros	Cons
			❇Expressive Language ❇Rigidity	❇Captures information beyond other tests (e.g., AQ; Kanne et al., 2012).	
Autism Quotient (AQ; Baron-Cohen et al., 2001)	Autism in individuals with average intelligence	Children, adolescents and adults	<u>Five factors</u>: ❇Social Skills ❇Attention Switching ❇Attention to Detail ❇Communication ❇Imagination	❇Acceptable internal consistency (Baron-Cohen et al., 2001) and strong test-retest reliability for total scores (Gravetter & Forzano, 2009). ❇AQ subscales recomposed in a three-factor model (DSM-IV; APA, 1994): Social skills, Details/patterns, and Communication/mindreading (English et al., 2019; Hurst et al., 2007; Russell-Smith et al., 2011). ❇Applied in the general population. ❇Commonly and worldwide applied to measure BAP. ❇A combination of relative and control samples suggests that AQ	❇15-20 min to complete. ❇Little knowledge of each subscale reliability (Zhang et al., 2016). ❇The only test about each subscale reliability for the original AQ counts on a very small sample (17; Baron-Cohen et al., 2001), remaining studies parts from different adapted versions. ❇An inconsistent number of factors (e.g., two-factor solutions, Hoekstra et al., 2008, 2011; Valla et al., 2010; three-factor solutions, Austin, 2005; Eriksson, 2013; Hurst et al., 2007; Ingersoll et al., 2011; Palmer et al., 2015; four-factor solutions, Saito et al., 2014; and five-factor solutions, Kloosterman et al., 2011), also through versions (e.g., two factors in AQ-Short version; Hoekstra et

Test	Target construct	Target population	Ground operationalization	Pros	Cons
				could serve to study the distribution of autistic traits (Ruzich et al., 2017). ❊Multiple test versions: self-report version, caregiver or parent-report version (for children/adolescents), short version (Hoekstra et al., 2011). ❊Translated versions (e.g., *Chinese*, Zhang et al., 2016; *French-Canadian*, Lepage et al., 2009; *German*, Freitag et al, 2007; *Italian*, Ruta et al., 2012; *Japanese*, Kurita et al., 2005).	al., 2011) and translations (e.g., three factors in the AQ-K, Freitag et al., 2007). ❊Varying scoring methods: dichotomous vs. binary scoring (Baron-Cohen et al., 2001; Ruzich et al., 2015) vs. Likert scoring, (Bishop et al., 2004; Wheelwright et al., 2010). ❊Possible gaps in some subthreshold ASD traits as originally emphasized Asperger traits (Kanne et al., 2012). ❊Does not share the two-domain structure of the ASD.
Social Responsiveness Scale—Second Edition (SRS-II (Constantino & Gruber, 2012)	Autism in individuals with average intelligence	2 years, 6 months through adulthood;	<u>Five subscales</u>: ❊Social Awareness ❊Social Cognition ❊Social Communication ❊Social Motivation ❊Restrictive and Repetitive Behaviors	❊Based on a sample of 1,906 individuals. ❊Total score, subscale scores, and scores for two DSM-V compatible subscales (Constantino & Gruber, 2012): Social Communication and Interaction, and Restricted	❊15-20 min to complete. ❊Qualification required. ❊Limited access (copyright by Western Psychological Services). ❊An inconsistent number of factors (e.g., one-factor solution, Constantino et al., 2000, 2004; Constantino & Todd,

Test	Target construct	Target population	Ground operationalization	Pros	Cons
				Interests and Repetitive Behaviours. ❈Facilitate the comparison of symptoms and determine whether a person may meet the most current diagnostic criteria for ASD. ❈Applied in the general population. ❈Commonly and worldwide applied to measure BAP. ❈Adequate subscales' reliability and validity with other autism subscales (ADOS, AQ, SCQ; Bölte, 2011; Bölte et al., 2008; Constantino & Todd, 2000, 2003; Virkud et al., 2008). ❈Multiple versions: Parent/teacher rating scale (from 2.5 to 18 years), self or informant-	2003), also in adapted versions (e.g., one factor in the *German* version, Bölte et al., 2008; five factors in the *Spanish* official version, Constantino & Gruber, 2017). ❈Though applied in community samples, it aims to diagnose adults with ASD (whether a person meets ASD criteria; Constantino & Gruber, 2012). ❈Does not share the two-domain structure of the ASD.

Test	Target construct	Target population	Ground operationalization	Pros	Cons
				report (adults up to 19), and short-forms. ❇Official Spanish translations of all its versions. ❇Adaptations to multiple languages (*German,* Bölte, 2011; *Vietnamese,* Nguyen et al., 2019)	
Social Communication Questionnaire (SCQ; Berument et al., 1999; Rutter et al., 2003)	Autism	High-risk children	<u>Three domains</u>: ❇Reciprocal Social Interaction ❇Communication ❇Restricted, Repetitive, and Stereotyped Patterns	❇Corresponds to the diagnosis of autism in DSM-IV (APA, 1994). ❇10 min application and 5 min evaluation. ❇Translated versions (*German,* Bölte et al., 2008, *Chinese,* Gau et al., 2011).	❇Dichotomous scoring impedes the dimensional approach and hinders the exploration of autistic traits in the general population. ❇Validation supports its use in suspected individuals rather than the general population. ❇Does not share the two-domain structure of the ASD.
Children's Communication Checklist—2 (CCC-2; Bishop, 2003)	Pragmatic Language Impairment	School-aged siblings and twins of diagnosed individuals	<u>Ten subscales</u>: ❇Speech ❇Syntax ❇Semantics ❇Interests ❇Coherence ❇Use of Context	❇Three composite scores (general communication, pragmatic, and social interaction deviance).	❇Not designed for measuring autism. ❇Does not share ASD internal structure. ❇BAP conceptualization is not clear (Ingersoll & Wainer, 2014).

Test	Target construct	Target population	Ground operationalization	Pros	Cons
			❊Inappropriate initiation ❊Stereotyped Language ❊Non-verbal Communication ❊Social Relations	❊Adapted versions (*Norwegian,* Helland et al., 2009; *Serbian*, Glumbic & Brojčin, 2012). ❊Reasonable internal consistency.	❊Only for children. ❊Inconsistencies in factor-solutions across versions (e.g., three factors in the Serbian version, Glumbic & Brojčin, 2012). ❊Does not share the two-domain structure of the ASD.
Childhood Asperger Syndrome Test (CAST; Scott et al., 2002)	Asperger Syndrome	School-aged siblings and twins of diagnosed individuals (unaffected); parent-report	Three factors: ❊Social Impairments ❊Communication Impairments ❊Repetitive or Stereotyped Behaviours	❊Three domains of impairment of ASC in DSM-IV-TR (Williams et al., 2005). ❊Parent or caregiver report ❊Dichotomous rating. ❊Includes special needs section.	❊Divergences in the CAST internal structure across studies in the same country (e.g., Williams, 2005, identified four factors: social behaviour and routines, speech and communication, peer relationships, and imaginative play) and adapted versions (two factors, Sun et al., 2014). ❊BAP conceptualization is not clear (Ingersoll & Wainer, 2014). ❊Targets Asperger traits. ❊Only for children.

Test	Target construct	Target population	Ground operationalization	Pros	Cons
					❇Does not share the two-domain structure of the ASD (Ingersoll & Wainer, 2014).
Social and Communication Disorders Checklist (SCDC; Skuse et al., 2005)	Autism in unaffected children	School-aged siblings and twins of diagnosed individuals	<u>Three factors</u>: ❇Social Impairments ❇Communication Impairments ❇Repetitive or Stereotyped Behaviours (reflecting the autistic triad) and three additional items for general behavioural problems.	❇Brief form. ❇Measures heritable characteristics. ❇Adequate internal consistency, test-retest reliability, and good discriminant validity (Skuse et al., 2005). ❇Parent report. ❇More discriminant in the general population.	❇BAP conceptualization is not clear (Ingersoll & Wainer, 2014). ❇Only for children. ❇Does not share the two-domain structure of the ASD.

Another alarming aspect is the fact that some measures have not shown even internal structure stability across studies with the same test version (e.g. the AQ has provided different factor solutions, ranging from two-factor, Hoesktra et al., 2011, up to five-factor solutions, Kloosterman et al., 2011; see Table 1). Accordingly, despite the virtues of the BAPQ (see Table 1), previous validity evidence based on the BAPQ internal structure has pointed to dimensionality problems in one of its subscales (i.e. the Pragmatic Language subscale; see e.g. Godoy-Giménez et al., 2018, Sasson et al., 2013). On the other hand, by including more than three factors, as is the case with other measurement tools, tests may also err on construct *irrelevance*, which refers to the degree to which test scores may be affected by extraneous processes to the test's intended purposes (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 2014; e.g. the Imagination factor from the AQ does not directly correspond to any subdomain of the ASD definition; see Table 1).

According to the *Standards* (AERA et al., 2014), whether two tests define the same construct differently and, thus, contain different internal structures, those tests are assessing two essentially different things. Consequently, the application of these tests, whether isolated or collectively, could drive observation of different symptoms, thus challenging the interpretation of the test scores, hindering the comparison of studies conducted with different tests and, finally, limiting all of the benefits mentioned above. This could be even more worrying considering that none of the mentioned – and most frequently used – measurement tools share the updated ASD or BAP structure (see Table 1).

These problems, arising from the disparity of BAP operationalisation or the use of measurement tools for different objectives than those for which they were intended, are connected with additional drawbacks that make it difficult to rely on the results of BAP measurement. As such, Godoy-Giménez et al. (2021a) provided validity evidence based on test content through expert judgment about the relevance of the items from three of these questionnaires – the BAPQ, the AQ and the SRS-II – and the representativeness problems of the BAP in those tests. These problems may derive from construct misrepresentation, which could lead to failure to capture important aspects of the BAP (e.g. the AQ does not include key deficits of restricted and repetitive behaviours; see Table 1). Notably, the same panel of experts agreed that all of the selected BAPQ items (which grasps core BAP factors according to the autistic dimensions in the DSM-IV; APA, 1994), was relevant for measuring the updated BAP structure but, at the same time, not sufficiently representative of the entire updated operationalisation of the construct. In this regard, they did not represent the subdomain of hyper- or hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment from the ASD definition (APA, 2013).

An additional problem related to BAP measurement involves the international adaptation of the BAP measures, which may be exacerbated by the fact that both BAP and ASD expressions could be affected by the culture of the test-taker (e.g. social acceptance and/or social support may differ, or some specific behaviours could present with different frequency according to the culture, like lack of eye contact; Volkmar et al., 2014). Accordingly, providing comparability of inferences made from test version scores rests on whether scores represent the same construct as those from the original version of the test (AERA et al., 2014). To explore cross-cultural BAP expressions, at least configurational invariance (i.e. the basic factor structure is invariant across countries) must be reached. This type of invariance lets us assume that two subgroups of people conceptualise the BAP in the same way (Brown, 2015; Meredith, 1993). Nevertheless, it is this metric invariance that will allow us to compare the obtained ratings of both populations, as it suggests that two populations may respond to the items in identical ways (Chakraborty, 2017). Although there exist multiple international adaptations of BAP measures (e.g. the BAPQ has been adapted to Chinese, Hindi, Meera et al., 2015; Shi et al., 2015; or Spanish, Godoy-Giménez et al., 2018; see Table 1), available BAP tests do not show structure stability (or even configurational invariance) with their translated versions (e.g. the SRS-II has been structured in one single factor in its German version, Bölte et al., 2008; and reported five factors in the Spanish version, Constantino & Gruber, 2017; see Table 1) and this hinders the comparison of scores cross-culturally.

In conclusion, it is necessary for an updated BAP test to be constructed upon a two-domain structure but that also aims to carefully develop translated versions (following former guidelines; International Test Commission [ITC], 2017) so as to try to maintain structural stability, which would then enable the comparison of test scores cross-culturally.

The Present Work

Structure of the Broad Autism Phenotype—International Test

Our main objective is to design, develop and pilot a test, the Broad Autism Phenotype—International Test (BAP-IT) in two countries simultaneously, Spain and the United Kingdom. The BAP-IT aims to be a scientific tool to measure the updating of the BAP operationalisation proposed by Godoy-Giménez et al. (2021a) and congruent with the ASD view of the DSM-5 (APA, 2013). The objective of the test is to characterise the spectrum of autism in a two-dimensional structure: deficits in the social communication and social interactions subscale (hereafter SOCIAL-BAP subscale) and a pattern of restricted and repetitive behaviours and interests subscale (hereafter RIRE-BAP subscale) at BAP levels of severity. Finally, we also intend that the BAP-IT represent an invariant tool that measures these two domains in two countries, enabling further cross-cultural comparisons.

To achieve these objectives, one pilot study in three broad community samples (two from Spain, ES1 and ES2, and one from the United Kingdom, EN) will be performed for the construction of both test versions. Common factor methods will be employed to obtain the simplest BAP-IT factor structure (two correlated dimensions with only primary loadings; Brown, 2015) with an adequate content representation of the seven BAP/ASD subdomains (see Table 1 in the Appendix 5) and to select simultaneously the items that will finally compose the test in two languages. Our main premise is that the resulting test will count on a unidimensional structure inside each domain of the updated BAP operationalisation (Godoy-Giménez et al., 2021a) coherent with the ASD definition in the DSM-5. Additionally, through multigroup analysis, we will test whether that simplest factor structure will result in invariance across groups. Specifically, both configural and metric invariances between the Spanish and British versions of the BAP-IT are sought. Our second main goal is that both kinds of invariance will be grasped and, so, further comparisons among BAP-IT versions could be possible. Once we have achieved the simplest and invariant factor structure for the BAP-IT, further psychometric properties will be studied, such as reliability and validity in relation to other variables.

Further Evidence Supporting the Interpretation of the BAP-IT Scores: Validity Evidence Based on BAP-IT Subscales' Relationships with Other Variables

Validation is a key stage in test development and evaluation (AERA et al., 2014). Thus, we try to provide validity evidence based on BAP-IT scores' relationships with other subscales that refer to facets related to autistic or BAP-like behaviours. Hence, BAP-IT subscale scores (i.e. the SOCIAL-BAP and RIRE-BAP subscales) are expected to be related to test scores that have been previously applied to assess the BAP. In this regard, the scores of the BAPQ subscales (Aloof, Pragmatic Language, and Rigid factors) are studied in relation to BAP-IT scores as the original test for measuring the BAP. We hypothesise that SOCIAL-BAP will have a high positive relation with Aloof from the BAPQ, because both are based on the social domain of the ASD from the DSM (criterion A; either DSM-IV or DSM-5). In the same way, as mentioned, the Pragmatic Language scale score from the BAPQ considers not only deficits in social communication (criterion A) but also deficits representing restricted and repetitive behaviours (criterion B; APA, 2013). Thus, this factor will have a moderate-to-high positive relation with the SOCIAL-BAP and a low-to-moderate positive one with the RIRE-BAP, which will in turn be moderate-to-highly positively related to the Rigid factor, because both are supposed to include (to a certain extent) conduct representing restricted and repetitive behaviours and interest domain (criterion B; APA, 2013).

In the same way, the SRS-II, a test for assessing autism in adults with standard intelligence (Constantino & Gruber, 2012), is another test commonly applied for measuring the BAP. Importantly, as we have used the BAP operationalisation proposed in Godoy-Giménez et al. (2021a), which extended the ASD to lesser levels of severity in autistic-like behaviours, and considering that the SRS-II includes subscales that are closely related to the ASD definition in the DSM-5 (APA, 2013), it could be supposed that the BAP-IT will be positively linked to all of the subscales of the SRS-II. That is, SOCIAL-BAP should be positively related to problems in expressive social communication (Social Communication subscale), interpreting social signs (Social Cognition subscale), perceiving social cues (Social Awareness subscale) and motivation to engage in social interpersonal behaviour (Social Motivation subscale). Stronger relations may be expected for Social Motivation or Social Communication. On the contrary, it might be expected that behaviours like extreme worry about changes and distress about unexpected alterations in daily routines and resistance to trying new things (Restricted and Repetitive Behaviours subscale; RRB) will be positively connected to the RIRE-BAP.

We will also explore further validity evidence based on BAP-IT relations to variables that refer to facets similar to those included in the BAP construct. Specifically, these focus on restricted and repetitive behaviours which have not been commonly studied in the literature. The Resistance to Change Scale (RCS; Oreg, 2003) assesses the "dispositional resistance to change", which denotes the inherent predisposition of some individuals to do fight against changes. RIRE-BAP should be more strongly related to a preference for routines and predictable activities (Routine Seeking subscale) and disapproval of alternative ideas, perspectives and methods (Cognitive Rigidity subscale) but will also be connected with focus on disadvantages derived from change (Short-term focus subscale) and perceived discomfort during uncontrollable changes (Emotional Reaction subscale).

If these empirical relationships match with theoretical expectations, these mentioned relations will also enable the comparison of BAP-IT validity evidence and inferences made on its subscales to validity evidence from precedent tests for the BAP. After all of the efforts to improve the representation and measurement of the BAP, we expect that the BAP-IT and its subscales, SOCIAL-BAP and RIRE-BAP, will be more coherently associated with BAP-related indicators than the subscales of other tests. In particular, the relationships of the subscale scores of the BAPQ (Aloof, Rigid and Pragmatic Language) will be explored together with those from BAP-IT (SOCIAL-BAP and RIRE-BAP), because this test has been acknowledged as the original test of the BAP (Hurley et al., 2007).

Finally, going another step forward, we will use empirical network models to explore the partial correlations between variables to compare the predictive capability of the BAP-IT and the BAPQ subscale scores. Empirical network analysis will help us to explore how the two-dimensional solution of the BAP-IT (Godoy-Giménez et al., 2021a) and its content representing the operationalisation aligned with the one of the ASD in the DSM-5 (APA, 2013) explain the relationships of its subscales scores with the rest of the variables implicated, in contrast to the three-factor solution of the BAPQ (see Hurley et al., 2007), in a manner more aligned with the original definition of BAP and the DSM-IV (APA, 1994).

Method

The development of the BAP-IT took several subsequent stages following the guidelines of the *Standards* (AERA et al., 2014) and the ITC (2017; see the procedure section). Once the candidate items were constructed in both languages, they were applied to three different samples (two from Spain and one from the United Kingdom) in a broad pilot study together with the Spanish or the English versions of the BAPQ. Eventually, once the internal structure matched the operational definition, the next step was investigating the evidence based on relations with other variables.

Participants

First Spanish Community Sample (ES1)

First, quota sampling was performed. Seven quotas by age and gender were calculated for a sample of 600 participants to ensure access to a representative sample regarding the age and sex of the Spanish population.[12] Second, incidental snowballing took place, as we counted on the help of undergraduates from the University of Almería (Spain) to have access to a larger sample and complete the quotas, which involved themselves, their relatives and acquaintances. Undergraduates received two course credits for their collaboration in the dissemination of the study. Finally, we completed the data collection by incidental sampling until reaching a final sample of 970 participants, with people captured from social networks and the Spanish associations of autism.

Second Spanish community sample (ES2)

Six months later, we collected a new sample of 460 participants. Sampling and collection procedures were similar to those in ES1.

[12] Percentages of men and women per quotas from the Spanish population in 2020 (Instituto Español de Estadística [INE], in English: Spanish Statistics Institute; 2020): 18–24 (5.35% men; 5.10% women); 25–31 (5.93% men; 5.87% women); 32–38 (7.37% men; 7.74% women); 39–45 (9.18% men; 8.96% women); 46–52 (8.60% men; 8.51% women); 53–59 (7.58% men; 7.77% women); and 60–66 (5.98% men; 6.37% women).

English community sample (EN)

We contacted an international company (https://www.gfk.com), which was in charge of recruiting a representative sample from the United Kingdom. This time, we restricted accessibility to the questionnaire by specifying the eight quotas (i.e. age and gender) on a sample of 500 participants (quota size was proportional to the corresponding population groups in England[13]), and we informed the company about the targeted sample composition. The company posted a notice with the questionnaire in an online panel for respondents, and participants who met the criteria (quotas by age and gender) selected the survey themselves (Self-Selected Sampling). A final group of 533 respondents participated in the study. They received €6 for their collaboration in the study.

Regarding inclusion criteria, participants from the four samples were asked about their nationality and how many years they had been living in Spain or the United Kingdom. Participants who completed the whole questionnaire very quickly (in less than 15 minutes) were excluded from the study and reinforcers were removed (46 participants were excluded and replaced by new ones). Other sociodemographic variables out of the scope of this study were also recorded (e.g. having a diagnosis of a mental disorder, belonging to a family with a history of autism or any related developmental disorders, medicalisation and psychiatric and/or psychological treatments).

Procedure: Stages in the Construction and Validation of the BAP-IT

Following Downing and Haladyna (2006), our protocol included: construct definition (see the introduction section), item generation, item selection and piloting of items.

Item Generation and Item Selection in Two Languages

The BAP-IT aims to be a short and handy self-report test for the international exploration of BAP expressions. The items should thus be consistent with these goals (Osterlind, 2002). The entire procedure for creating and selecting the items

[13] Percentages of men and women per quotas from the English population in 2020 (Office for National Statistics, 2020): 18–24 (5.99% men; 5.66% women); 25–31 (6.60% men; 6.45% women); 32–38 (6.33% men; 6.44% women); 39–45 (5.86% men; 5.95% women); 46–52 (6.47% men; 6.66% women); 53–59 (6.37% men; 6.58% women); 60–66 (5.17% men; 5.41% women); and 67–72 (0.58% men; 0.64% women).

was established through a careful protocol (see AERA et al., 2014; Irwin et al., 2018; Lane et al., 2015). Items used a Likert-type rating scale and asked about the frequency of specific behaviours and personality preferences (Haladyna et al., 2002). Similarly, they targeted either Spanish or English adults from the general population with standard intelligence and reading ability.

First, a representative sample of BAP key expressions (352 items for SOCIAL-BAP and 246 for RIRE-BAP) was produced in Spanish to measure the BAP expressions according to the updated BAP operationalisation presented in the introduction section (see also Godoy-Giménez et al., 2021a). Then, candidate items were constructed by an expert in psychometrics (PSJ attended the formal aspects) and one in autism (MGG focused on content). Item generation was theory driven and carefully developed considering both the formal aspects of the items (the *Standards*; AERA et al., 2014; Lane et al., 2015; Osterlind, 2002) and the item content, which was based on the DSM-5 diagnostic criteria for ASD (APA, 2013) as well as the contributions of authors of reference in the field of autism and BAP (e.g. Lord et al., 2018, 2020; Piven et al., 1997a, b).

For item selection, the first step was to review the formal aspects of the items (598 items in total; Irwin et al., 2018), as they may affect the way the construct is measured on the scale (Haynes et al., 1995). Four of the members of our research lab (AGR, AGP, FC, AFE) conducted an itemmetric assessment (Angleitner et al., 1986) after receiving training from our expert in psychometrics. The procedure consisted of rating independently from 1 to 4 (for example, 1 = *not clear,* 2 = *somewhat clear,* 3 = *quite clear* and 4 = *very clear*) each item according to the following properties: clarity, comprehensibility, concreteness, degree of self-reference and evaluation of the items (Angleitner et al., 1986). Only items that showed expert agreement (i.e. with averages among the reviewers ≥ 3.8) on all of the assessed variables were incorporated into the pool of items that would be tested empirically (275 items for SOCIAL-BAP and 229 for RIRE-BAP).

As all of the autistic expressions seemed relevant for BAP diagnosis (APA, 2013), a subsequent content evaluation (Grant & Davis, 1997) assessed the relevance of each item for its specific subdomain (always choosing the items with the greatest content-based relationship with the subdomain), content redundancy and the representativeness of the BAP domains and subdomains. Two members of our research lab (FC and AFE) rated the items using a Likert-type ordinal scale with four possible responses (e.g. 1 = *not relevant*, 2 = *somewhat relevant*, 3 = *quite relevant* and 4 = *very relevant*). According to the content validity index (CVI; Hambleton et al., 1978; Waltz & Bausell, 1983), only the items that received a rating of 3 or 4 from both experts were maintained. Final decisions on items (i.e. conserving or deleting) were based on both data analysis and comments from the experts. Finally, 242 Spanish-based items were proposed for translation into English as possible candidates of BAP indicators.

The 242 items were adapted to English following the ITC guidelines (2017; www.InTestCom.org). First, two bilingual translators (one native English and one native Spanish, both bilinguals in both English and Spanish) conducted two separate forward translations of the items to English and both works were subsequently unified. The items were then translated back into Spanish by one lab collaborator with an advanced level of English and Spanish (the same procedure was taken for translating the rating scale, test instructions and so on). Afterwards, two bilingual associate professors in the Psychology Degree Programme of Bangor University, Wales and the Psychology College of Dublin, Ireland, who collaborated with the research, evaluated whether the items were targeting the correct BAP indicators in both languages. Finally, the six authors of the study revised the resulting pool of items, ensuring that it maintained the representation of all of the BAP/ASD subdomains and their balanced distribution (items assessing different severity expressions and balanced weight of the subdomains in the final number of items). For this, the number of items inside each subdomain was controlled and items that expressed similar content (e.g. items that targeted the same BAP deficit but expressed examples in different contexts) were eliminated, resulting in 158 final items (see Table 1 in the Appendix 5).

Piloting Stage

The data were compiled in a single online session which lasted approximately 30 minutes for each sample. Participants had access to the questionnaire after following a link, and it was completely anonymous. Questionnaires should be completed on large tablets, laptops and/or desktop computers to ensure the correct presentation of the items. All of the participants read the instructions and the data treatment ethics and gave their consent to participate in the study and to the use of their data only for scientific purposes before starting with the questionnaire. They were asked to read each question carefully and answer them sincerely. They were also warned that aleatory patterns of responses would be tracked and together with very short time lapses were sufficient reason to remove collaboration reinforcers.

The candidate items for the BAP-IT along with an alternative pool of items (see the Instruments section) were always presented first, followed by the BAPQ. All of the participants finished the session by answering some socio-demographic questions (e.g. gender, age, educational level, nationality). All questions except some sociodemographic ones (i.e. those including sensitive or private content) were compulsory. Participants could not leave the session to retake it later and/or redo the questionnaire. Although they knew what the aim of the survey was (to validate a new questionnaire assessing the BAP), the names of the questionnaires included were not mentioned to prevent biased information.

Additional Validity Evidence

To gather validity evidence based on test relations to other variables, some additional tests were included after the BAPQ in the ES2 data collection. The ES2 group filled the BAP-IT candidate pool of items, the alternative pool of items, the BAPQ, the SRS-II, the RCS (in this order), and seven more questionnaires beyond the scope of the present study.

Instruments

BAP-IT Candidate Items

A pool of 158 items, in Spanish and English, was proposed as candidate items for the BAP-IT (see procedure). Items were distributed in two domains and were coded regarding the domain and specific subdomain and key behaviour[14] they represent (e.g. SA1A7: candidate item for the domain SOCIAL-BAP, *S*, subdomain deficits in

[14] SA, deficits in social-emotional reciprocity; SB, deficits in nonverbal communicative behaviours used for social interaction; SC, deficits in developing, maintaining, and understanding relationships; RA, stereotyped or repetitive motor movements, use of objects or speech; RB, insistence on sameness, inflexible adherence to routines, or ritualised patterns or verbal nonverbal behaviour; RC, highly restricted, fixated interests that are abnormal in intensity or focus; RD, hyper- or hypo-reactivity to sensory input or unusual interests in sensory aspects of the environment (APA, 2013).

social-emotional reciprocity, *A1*, and the seventh example of Abnormal social approach and social communication interferences, *A7*). Appendix 5 presents the candidate pool of items (see Table 1) and the domains and subdomain to which they belong (see Table 1 and Table 3, which include the domains and subdomains to which the items belonged). Items were rated regarding the frequency of the behaviour they were expressing on a six-point Likert-scale (1 = *never*, 2 = *almost never*, 3 = *occasionally,* 4 = *often,* 5 = *almost always and* 6 = *always*). Two overall scores were intended, one for SOCIAL-BAP and one for RIRE-BAP; the higher the score, the more severe the expressions of the behaviours that the dimension contains.

The final items, the test instructions and the rating scale are displayed in Table 5. We used the instructions and rating scales presented in Godoy-Giménez et al. (2021a; Appendix 4. Document 2), and the same sample of experts assessed the degree of adequacy of test instructions and rating scale to the measurement objective of the test and target population through expert judgment (n = 20; for the composition of the sample of experts, see Godoy-Giménez et al., 2021a). Three experts considered that the rating scale could be improved (15%), ten experts considered it was adequate (50%), six said that it was very adequate (30%) and one response was missing. In sum, 80% of the experts agreed with the adequacy of the rating scale. Regarding test instructions, one expert considered that they could be improved (5%), nine felt they were adequate (45%), nine others rated them as very adequate (45%) and one response was missing. In total, 90% of the experts expressed their agreement with the instructions of the test.

Alternative Pool of Items

In a previous study (Godoy-Giménez et al., 2021a), through expert judgment, it was tested whether the items of three worldwide applied and adapted tests, the AQ (Baron-Cohen et al., 2001), the SRS-II (Constantino & Gruber, 2012) and the BAPQ (Hurley et al., 2007), could be reshaped in the updated BAP conceptualisation of two main domains. From this study, 48 items from the mentioned tests were selected as they could have adequate content for assessing the BAP according to its updated definition (Godoy-Giménez et al., 2021a). We decided to run an empirical study (EFA in the CFA framework —E/CFA; Brown, 2015; Jöreskog & Sörbom, 1979;

Muthén & Muthén, 1998–2012) to select which of those 48 items conformed to a two-dimensional structure according to the BAP updated definition (the whole procedure and its results can be consulted in the Appendix 6. Document 1). Ten items with high loading in their respective domains (above .40) and without cross-loadings (below .30) were selected (SOCIAL-BAP counted on seven items and RIRE-BAP on three).

We took selected items and applied them together with the *BAP-IT Candidate Items*. The inclusion of this *Alternative Pool of Items* was motivated by both the empirical and theoretical evidence of their relevance for assessing the BAP around two domains (Godoy-Giménez et al., 2021a). The alternative pool of items is included in the Appendix 5 (see Table 2; items sorted by subdomain, code and Spanish and English content) as well as the domains and subdomain to which the items belonged (see Table 3).

The BAPQ

The BAPQ (Hurley et al., 2007) is a 36-item screening questionnaire specifically designed to assess the BAP in relatives of people diagnosed with ASD. BAPQ items are grouped into three subscales (Hurley et al., 2007; Piven et al., 1997a, b): aloofness, rigidity and pragmatic language problems. The correspondence of the items and the subscales as well as the reversed items can be found in Hurley et al. (2007). Respondents express their agreement with each item on a 6-point Likert scale (1 = *very rarely*; 6 = *very often*). The higher the score, the more severe the deficit is. The BAPQ-SP (Godoy-Giménez et al., 2018; Godoy-Giménez, 2021b [manuscript submitted for publication]) corresponds to the official Spanish adaptation of the test.

The Spanish version of the SRS-II (adult self-report version)

The SRS-II (Constantino & Gruber, 2017) is a 65-item self-report test examining the severity of autistic behaviours in adults with normal intelligence. Items are distributed in five subscales: Social Awareness, Social Cognition, Social Communication, Social Motivation and Restricted Interests and Repetitive Behaviours; items from each subscale and reversed items can be consulted in Constantino and Gruber (2012; 2017). Each item in the test is rated on a 5-point Likert scale (1 = *not true*; 5 = *almost always true*). The higher the score, the more severe the deficit is. The Spanish adult self-report version of the SRS-II was developed by Constantino and Gruber (2017).

The Spanish version of the RCS

The RCS (Oreg, 2003) contains 17 items that assess dispositional resistance to change in four dimensions: Emotional Reaction, Short-term Focus, Cognitive Rigidity and Routine Seeking. Respondents express their agreement with each item on a 6-point Likert scale (1 = *strongly disagree*; 6 = *strongly agree*), and a composite RTC score is provided. The higher the score, the greater the dispositional resistance to change. Here we applied the Spanish version of the RCS, which was presented in Arciniega and González (2009).

Compliance with Ethical Standards

This study received ethics approval from the local Human Research Ethics Committees of the university involved in the recruitment process, which was conducted under the approved guidelines and the Declaration of Helsinki. Written informed consent was obtained from all participants before conducting the study.

Results

The development of the BAP-IT took several refining steps to find the most appropriate content according to its planned/theoretical structure and operationalisation. The rationale of the analysis (depicted in Table 2) together with the specifications of the models are presented at the beginning of each section. Justification of the sample order in the analysis relies on sample temporal disposition. The analysis started with the first and larger sample, ES1, and continued by combining it with the EN sample (multi-group analysis), as our first objective was to obtain an invariant test structure. Finally, the ES2 was incorporated into the study to select additional items.

Table 2

Rationale of the Analysis

Steps	N cases	Sample group	Analysis/actions performed	Models
Scale Construction and Validity Evidence Based on the Internal Structure				
Step.1 Exploratory Item Inspection According Unidimensionality	970	ES1	Separated EFAs in SOCIAL-BAP and RIRE-BAP	
Step.2.a. Initial Simple Factor Structure Item Selection in ES1	970	ES1	E/CFAs	Step.2.a.Model1 Step.2.a.Model1$_{\text{Fitted}}$
Step.2.b. Item Selection According to Multiple Group Simple Factor Structure and Language Invariance	1503	ES1 EN	Multigroup CFAs for configural and metric invariance	Step.2.b.Model2 Step.2.b.Model3$_{\text{MINV}}$ Step.2.b.Model3 $_{\text{PMINV}}$ Step.2.b.Model3$_{\text{FREE}}$
Step.2.c. Content Reconstruction, Selection of Additional Items to Fill Content Gaps, Structure and Language Invariance, Structural Invariance (equal correlations)	1963	ES1 ES2 EN	Empirical Network Analysis for content reconstruction Multigroup CFAs for configural and metric invariance Multigroup CFAs for studying equality in lantent variables correlation among groups	Step.2.c.Model1 Step.2.c.Model2 Step.2.c.Model3$_{\text{MINV}}$ Step.2.c.Model3$_{\text{PMINV}}$ Step.2.c.Model3$_{\text{FREE}}$ Step.2.c.Model4
Step.3. SOCIAL-BAP and RIRE-BAP scores' reliability	1963	ES1 ES2 EN	Omega estimations	Step2.c.Model4
Validity Evidence Based on the BAP-IT Associations with Related Subscales				
Step.1. BAP-IT Relations with the BAPQ Subscales	1963	ES1 ES2 EN	Relationships among SOCIAL-BAP, RIRE-BAP, Aloof, Rigid, and Pragmatic Language	Step2.c.Model4
Step.2. Further Validity Evidence Based on BAP-IT Relations with BAPQ, RCS and SRS-II subscales in ES2.	460	ES2	Relationships among SOCIAL-BAP, RIRE-BAP, Aloof, Rigid, Pragmatic Language, Social Motivation, Social Communication, Social Cognition, Social Awareness, Restricted and Repetitive Behaviours (RBB), Cognitive Rigidity, Emotional Reaction, Routine Seeking and Shor-term Focus.	Step2.c.Model4

Scale Construction and Validity Evidence Based on Internal Structure

We drew from an exploratory item inspection towards a confirmatory framework in search of the simplest and unidimensional (within each domain: SOCIAL-BAP and RIRE-BAP) BAP-IT model. Descriptive statistics and multivariate normality of all of the BAP-IT and BAPQ indicators were analysed with Mardia's multivariate skewness and kurtosis (Mardia, 1970) with the MVN package 5.8 version (2019) for assessing multivariate normality in R software, version 4.0.3 (R Core Team, 2020). All values were statistically significant (p < .01; see Table 5 Appendix 5). Results from the descriptive analysis are displayed in Tables 4 and 5 in the Appendix 5 (*BAP-IT Candidate Items*, *Alternative Pool of Items, and BAPQ subscales*).

Step 1: Exploratory Item Analysis According to Unidimensionality (ES1 sample)

Our first step was to perform one independent EFA in each domain (one for SOCIAL-BAP and the other for RIRE-BAP) to select those items with strong loadings in the primary factor (above .30) and without cross-loadings on secondary factors (below .30) to prevent multidimensionality and the possible presence of local dependency (on those with more than one item loading on secondary factors). We started with the *BAP-IT Candidate Items* for the ES1 group. EFAs were conducted with principal axis factoring (due to the lack of multi-normality) in SPSS Statistics 26 (IBM Corp, 2020). The number of factors to extract in each solution was determined using exploratory graph analysis (Golino & Epskamp, 2017; eleven factors for SOCIAL-BAP accounted for a total of 61.01% variance and 17 for RIRE-BAP for a total of 61.51% variance). Factor loadings of both EFAs (the unrotated initial factor loading matrix was analysed) and the final selected items (35 for SOCIAL-BAP and 46 for RIRE-BAP) can be seen in Tables 6 and 7 in the Appendix 5. The selected items represented the content of the seven BAP/ASD subdomains.

Step 2: Multiple Group Confirmatory Factor Analysess According to Simple Factor Structure and Language Invariance

General specifications in all of the models in the CFA framework are presented below. As the descriptive results showed that the data in the different subgroups did not follow a normal distribution, the model estimator was *maximum likelihood estimation with robust standard errors* (MLR) due to its robustness to non-normality, ability to accommodate missing data and computational efficiency (Muthén & Asparouhov, 2011). In the EFA in the CFA framework, factor variances were fixed to one, factor means to zero and factor covariance was freely estimated. Cross-loadings were fixed to zero for anchor items (RD3A14 was the anchor item for SOCIAL-BAP and SC4A5 for RIRE-BAP) and freed for the rest of the items. In the CFAs, first loadings were fixed to 1.0 in all of the models to set the metric of the SOCIAL-BAP and RIRE-BAP factors. Acceptable model fit was established at RMSEA (≤ 0.06), CFI (≥ 0.95) and TLI (≥ 0.95) following Hu and Bentler (1999). When RMSEA is higher than or equal to .10 and CFI values are lower than .90, the model should be discarded (Brown, 2015). The specification search (Leamer, 1978) of localised areas of misfit was tested with modification indices (MIs; MacCallum, 1986; Sörbom, 1989). The strategy for conducting the specification search was as described by MacCallum (1986) for both item selection and metric invariance. Regarding item selection, we first selected the items with attention to MIs concerning cross-loadings and, afterwards, correlated uniquenesses, deleting, one-by-one, items that presented MIs > 10. Afterwards, those items presenting a strong lack of metric invariance were also deleted. However, as most autistic behavioural expressions may be susceptible to many cultural idiosyncrasies, when considered necessary (as in the case of maintaining central contents of the BAP while preserving the two-dimensional structure), achieving partial metric invariance was intended. Further considerations of the impact of cultural differences on these specific BAP expressions will be theoretically argued in the discussion section. Moreover, in this selection procedure, we tried to be loyal to the BAP-IT operationalisation, so the whole strategy was conditioned upon the maintenance of an adequate content representation of all seven subdomains in their respective domains. Comparisons of models were accomplished using Satorra-Bentler scaled chi-square (2001) to calculate chi-square difference testing using the MLR estimator. Analyses were performed in MPlus v7.11. (Muthén & Muthén, 2012). Table 8 in the Appendix 5 contains the retained and discarded items on each step.

Step 2a. Initial Simple Factor Structure Item Selection in ES1 sample.

An EFA in the CFA framework (Brown, 2015; Jöreskog & Sörbom, 1979) was applied to the 81 selected *BAP-IT Candidate Items* in the ES1 sample. The data set contained 970 complete cases. We drew from the initial model (Step2.a.Model1; its fit indices can be seen in Table 3) and selected those items with factor loadings above .30 in their respective factors and with weak or null loadings on the other factor/domain (below .20). Nineteen items were discarded due to cross-loadings. Furthermore, in this preliminary selection, in the Spanish language, the items indicating local dependency, with MIs above 20, were discarded (25 items). From a total of 81 items, 37 final ones were selected (16 for SOCIAL-BAP and 21 for RIRE-BAP; selected and discarded items can be found in Table 8 in the Appendix 5) resulting in the Step2.a.Model1$_{\text{Fitted}}$ model (fit indices are presented in Table 3) with a correlation between latent variables equal to .45 (p < .001). Factor loadings of the Step2.a.Model1$_{\text{Fitted}}$ for SOCIAL-BAP and RIRE-BAP are displayed in Appendix 5, Table 9.

Table 3

Model Fits

BAP-IT two-factor models	N items	χ2(df)	χ2 contribution from each group	Scaling Correct. Factor	RMSEA [90% CI]	CFI	TLI	CD	TRd	∆df	p-value
Step.2.a. Initial Simple Factor Structure Item Selection in ES1											
Step.2.a.Model1	81	9140.338 (308)		1.18	.05[.04, .05]	.79	.78				
Step.2.a.Model1$_{\text{Fitted}}$	37	1550.770(593)		1.18	.04[.04, .04]	.91	.89				
Step.2.b. Item Selection According to Multiple Group Simple Factor Structure and Language Invariance											
Step.2.b.Model2	19	605.74(302)	ES1 (323.21) EN (282.53)	1.23	.04[.03, .04]	.95	.95				
Step.2.b.Model3$_{\text{FREE}}$	17	464.47(236)	ES1 (265.05) EN (199.42)	1.24	.04[.03, .04]	.96	.95				
Step.2.b.Model3$_{\text{MINV}}$ (Contrast with Step.2.b.Model3$_{\text{FREE}}$)	17	490.76(251)	ES1 (276.86) EN (213.89)	1.22	.04 [.03, .04]	.96	.95	.91	25.17	15.00	.04
Step.2.b.Model3$_{\text{PMINV}}$ (Contrast with Step.2.b.Model3$_{\text{FREE}}$)	17	484.90(250)	ES1 (273.63) EN (211.27)	1.22	.04[0.3, .04]	.96	.96	11.18	1419.69	14.00	.14
Step.2.c.. Content Reconstruction, Selection of Additional Items to Fill Content Gaps											
Step.2.c.Model1	26	2346.644(894)	ES1 (968.40) ES2 (622.41) EN (755.83)	1.22	.05 [.05, .05]	.88	.87				
Step2.c.Model2	22	1367.78(624)	ES1 (559.72) ES2 (348.37) EN (459.69)	1.23	.05[.04, .05]	.93	.92				
Step.2.c.Model3$_{\text{FREE}}$	20	1175.67(507)	ES1 (504.02) ES2 (285.26) EN (386.38)	1.24	.05[.04, .05]	.93	.92				

BAP-IT two-factor models	**N items**	**χ2(df)**	**χ2 contribution from each group**	**Scaling Correct. Factor**	**RMSEA [90% CI]**	**CFI**	**TLI**	**CD**	**TRd**	**△df**	**p-value**
Step.2.c.Model3$_{MINV}$ (Contrast with Step.2.c.Model3$_{FREE}$)	20	1265.291(543)	ES1 (531.31) ES2 (305.51) EN (428.48)	1.24	.05[.04, .05]	.92	.92	1.24	89.62	36.00	.00
Step.2.c.Model3$_{PMINV}$ (Contrast with Step.2.c.Model3$_{FREE}$)	20	1216.85(539)	ES1 (516.73) ES2 (299.84) EN (400.28)	1.24	.04[.04, .05]	.92	.92	1.13	37.98	32.00	.22
Step.2.c.Model4 (Contrast with Step.2.c.Model3$_{PMINV}$)	20	1217.821(540)	ES1 (517.10) ES2 (300.43) EN (400.30)	1.24	.04[.04, .05]	.92	.92	1.24	.97	1.00	.32
Subsample 1:											
Step.2.c.Model3$_{PMINV}$	20	1374.55(701)	ES1 (374.62) ES2 (193.69) EN (448.60)	1.18	.05 [.05, .06]	.90	.90				
Step.2.c.Model4 (Contrast with Step.2.c.Model3$_{PMINV}$)	20	1389.09(711)	ES1 (553.60) ES2 (383.23) EN (551.33)	1.18	.05 [.05, .06]	.90	.90	1.18	14.53	10.00	.15
Subsample 2:											
Step.2.c.Model3$_{PMINV}$	20	1330.39(701)	ES1 (554.76) ES2 (341.18) EN (434.45)	1.18	.05 [.05, .06]	.91	.90				
Step.2.c.Model4 (Contrast with Step.2.c.Model3$_{PMINV}$)	20	1341.87(711)	ES1 (556.51) ES2 (346.73) EN (438.63)	1.18	.05 [.05, .06]	.91	.90	1.18	11.49	10.00	.32

Note. χ2(df): chi-square (degrees of freedom); RMSEA [90% CI]: root mean square error of approximation [90% confidence interval]; CFI: comparative fit index; TLI: Tucker–Lewis index; CD: difference test scaling correction; △df: difference is degrees of freedom; TRd: Sattora-Bentler Scaled chi-square difference.

Step 2b. Item Selection According to Multiple Group Simple Factor Structure and Language Invariance (ES1 and EN Samples).

Once a simple factor structure (coherent with the operational definition) was achieved in the ES1 sample, the next analysis aimed to select those items ensuring the same measurement model (i.e. configural invariance: two related factors with items of both domains without strong cross-loadings) with the same factor loadings (metric invariance) in the Spanish and English language versions. The codes of these items are presented in Appendix 5 Table 1 and the fit indices in Table 3 of the present document. The data set contained 1,503 cases from the ES1 group (n = 970) and the EN group (n = 533). There was one missing data item from one participant from the EN group. The item RD1A5 was involuntarily omitted in some questionnaires, and 66 further responses were missing in the EN group. Full information maximum likelihood was used to estimate the missing values (Muthén & Asparouhov, 2011).

We started with Step2.b.Model1 (that is, multi-group Step.2.a.Model1$_{\text{Fitted}}$). Inadequate fit values suggested that the model should be revised (latent variable correlations were .47, $p < .001$ for the ES1 group and .65, $p < .001$ for EN). Thus, item selection was performed as described above until a satisfactory solution was achieved. Step.2.b.Model2, with configural invariance among groups (fit indices are presented in Table 3), counted on 10 SOCIAL-BAP items and 9 RIRE-BAP items (selected and discarded items can be found in Table 8 in the Appendix 5; item codes are displayed in Table 1 of the Appendix 5).

Subsequently, the factor loadings of Step.2.b.Model2 were constrained equally among groups to test metric invariance (Step.2.b.Model2$_{\text{MINV}}$). Two items (SA3A16i, RA1A2) which showed a strong lack of invariance between ES1 and EN samples were removed, resulting in Step.2.b.Model3 (see Table 8 in the Appendix 5). We then tested the metric invariance in this model (Step.2.b.Model3$_{\text{MINV}}$; its fit indices are presented in Table 3), and the results indicated a lack of invariance in one item (the contrast with Step.2.b.Model3 can be seen in Table 3). Thus, we proceeded to achieve partial metric invariance (Step.2.b.Model3$_{\text{PMINV}}$; Byrne et al., 1989) so as to avoid missing the content of one subdomain. Specifically, the RC2B5 factor loading was freed, as we considered the content of the item highly relevant for the BAP operationalisation. Step.2.b.Model3$_{\text{PMINV}}$ (fit indices are in Table 3 and item loadings and intercepts in Table 4), with all but one item loading, constrained equally between samples, resulted in being statistically equivalent to Step.2.b.Model3$_{FREE}$ (the contrast between models is in Table 3). Latent variable correlations were .49 ($p < .001$) for ES1 group and .65 ($p < .001$) for EN. The average standardised factor loading for SOCIAL-BAP was M = .53 for both groups (metrically invariant), while for RIRE-BAP these were M = .51 in ES1 and M = .63 in EN.

Table 4

Factor Loadings for the Final Model

Group	Factor	Item	Standardised Estimate loading [non-standardized]	Intercepts
ES1	SOCIAL-BAP	SA1A7	.70[1.00]	2.82
		SA2A2	.61[0.91]	2.63
		SA3A7	.60[0.90]	3.34
		SA3A8	.66[1.07]	3.13
		SA3A11	.57[0.84]	2.89
		SB4B1	.60[0.93]	2.44
		ASRS60	.65[1.02]	2.49
		SC3B5	.65[0.98]	2.69
		SC4D1	.76[1.29]	3.32
	RIRE-BAP	RA2A10	.56[1.00]	2.28
		RA3A3	.51[1.05]	2.56
		RA3A6	.59[1.00]	2.21
		RB3B2	.38[0.80]	3.76
		RB4A2	.39[0.81]	3.06
		RC2A3	.45[0.92]	3.20
		RC2B5	.55[1.15]	2.62
		RC2C3	.56[1.11]	3.32
		RD1A5	.38[0.69]	2.37
		RD3A7	.48[0.82]	2.19
		RD3A14	.62[1.11]	2.11
ES2	SOCIAL-BAP	SA1A7	.65	2.58
		SA2A2	.58	2.49
		SA3A7	.55	2.92
		SA3A8	.61	2.69
		SA3A11	.56	2.57
		SB4B1	.63	2.10
		ASRS60	.61	2.21
		SC3B5	.64	2.30
		SC4D1	.75	2.81
	RIRE-BAP	RA2A10	.53	2.24
		RA3A3	.52	2.22
		RA3A6	.63	1.82
		RB3B2	.35	3.75
		RB4A2	.35	3.27

Group	Factor	Item	Standardised Estimate loading [non-standardized]	Intercepts
		RC2A3	.42	2.77
		RC2B5	.53	2.46
		RC2C3	.54	2.97
		RD1A5	.35	2.20
		RD3A7	.45	2.20
		RD3A14	.57	2.18
EN	SOCIAL-BAP	SA1A7	.69	2.58
		SA2A2	.59	2.83
		SA3A7	.64	3.45
		SA3A8	.64	3.25
		SA3A11	.61	3.18
		SB4B1	.79[1.24]	2.75
		ASRS60	.69	2.72
		SC3B5	.65	2.61
		SC4D1	.76	2.62
	RIRE-BAP	RA2A10	.69	3.20
		RA3A3	.74	2.48
		RA3A6	.75	2.38
		RB3B2	.51	2.20
		RB4A2	.61	3.18
		RC2A3	.44[0.59]	2.72
		RC2B5	.58[0.86]	2.84
		RC2C3	.55[0.80]	2.54
		RD1A5	.50	2.88
		RD3A7	.59	2.48
		RD3A14	.76	2.39

Note. Boldface indicates non-invariant items between ES groups and EN one.

Step 2c. Content Reconstruction, Selection of Additional Items to Fill Content Gaps (ES1, EN and ES2 samples).

The Step.2.b.Model3$_{PMINV}$ resulted in the SB subdomain from SOCIAL-BAP being without items representing their content and in the RB and RC from RIRE-BAP having only one item. We thus decided to include additional items regarding the under-represented subdomains. To decide which items should be reinserted and included in the following analysis, we recovered some items from the SB, RB and RC subdomains from the original pool of 158 *BAP-IT Candidate Items*. Specifically, besides item content (to fill content gaps without including redundant content), criteria recovering considered the factor matrix from Step 2a., those items of the mentioned subdomains which have presented factor loadings higher than .3 in their factor but, this time, without considering their secondary factor loadings. From all of the possible options, items with the highest factor loadings and lowest loadings in secondary dimensions and without redundant content were reinserted in the model. Additionally, items from the *Alternative Pool of Items* that belonged to SB, RB or RC were also selected as candidate items for content reconstruction in this step. These new candidate items can be found in Appendix 5 Table 8.

We aimed to select only those items with a coherent partial correlation (i.e. according to their content) with the SOCIAL-BAP and RIRE-BAP factor scores of Step.2.b.Model3$_{PMINV}$ while controlling for strong individual relationships between items (a possible manifestation of future local dependency; internal consistency of SOCIAL-BAP was .84 and RIRE-BAP was .78). We therefore decided to employ empirical network models to complete this step. Because partial invariance was achieved, we estimated the SOCIAL-BAP and RIRE-BAP factor scores (with the regression method model; Asparouhov & Muthén, 2010), in ES1 and EN samples, in Step.2.b.Model3$_{PMINV}$, and we merged them into one data set (to gain statistical power) including all of the newly reintegrated items (see Table 8 in the Appendix 5). We then displayed an empirical network using the R-package *bootnet* (Epskamp et al., 2017) applying the gLasso algorithm (a variant of the LASSO regularisation techniques; Friedman et al., 2014) and the Extended Bayesian Information Criterium (EBIC, Chen & Chen, 2009) to choose the tuning parameter (with the ES1 and EN samples merged). The edges were refitted without LASSO regularisation to avoid edge weights-estimations “biased toward zero due to shrinkage” (Epskamp & Fried, 2018, p.12). The adjacency matrix is displayed in Table 10, and the initial and final networks can be seen in Figures 1a and 1b (all in the Appendix 5). Inside each subdomain, items representing a partial correlation with the expected factor score and lesser correlation with the other domains were selected (to avoid ulterior cross-loadings). Likewise, selection criteria also took into account that items did not show high partial correlations with other items (to avoid local dependency). Finally, at least three items from each subdomain should remain. The selected items were

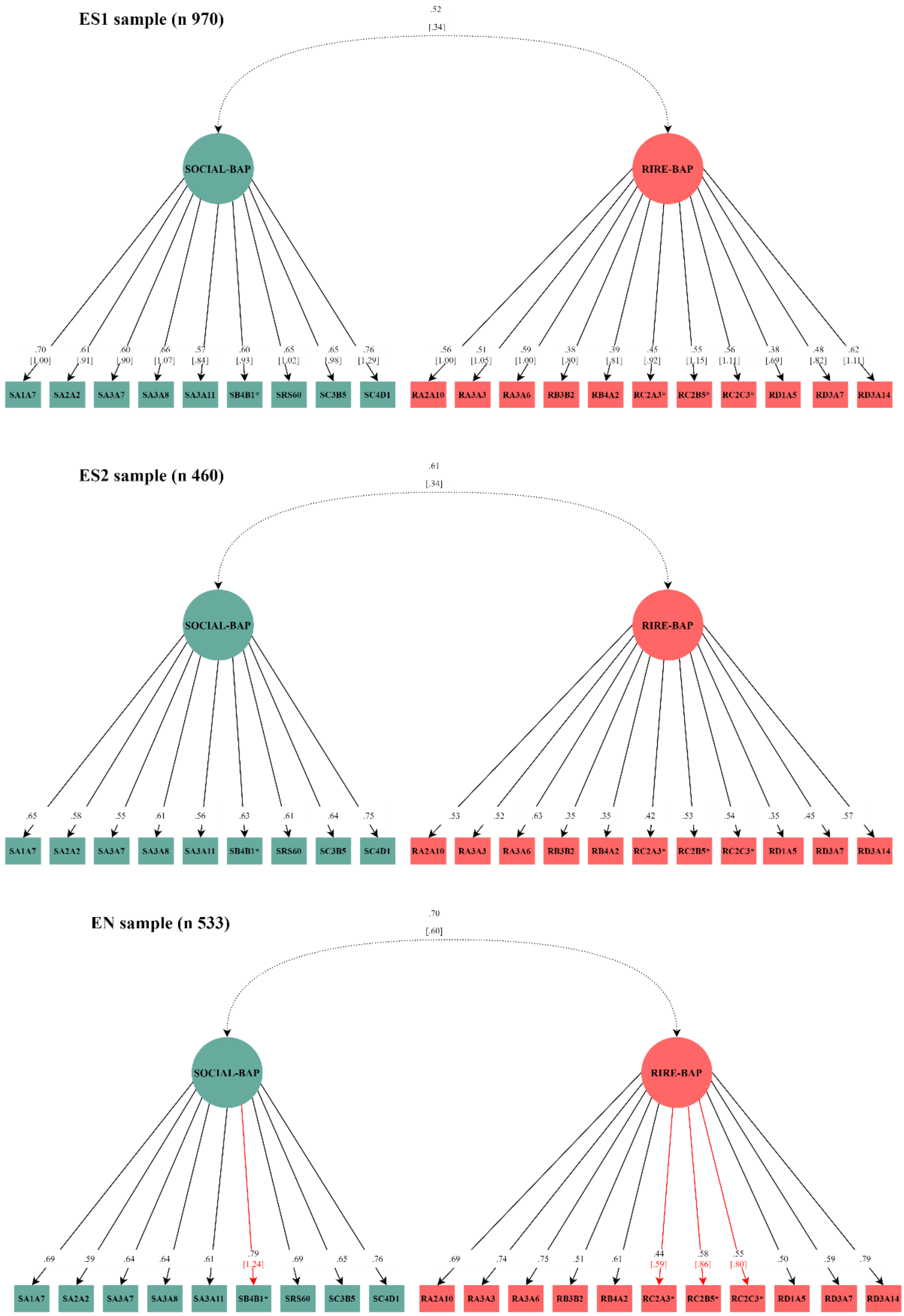
SB2B1i, SB4B1, SRS60 (SB), RB4A2, RB4A13, and RB5A3 (RB), and RC2A3, RC2C3, and SRS31 (RC; Table 8 in Appendix 5).

Once the new items were selected, we followed the approach presented in Step 2b but added a new data set to the multigroup model. The final data set contained 1963 complete cases from ES1 (n 970), ES2 (n 460), and EN (n 533) groups. Thus, the multigroup Step.2.c.Model1 included 26 items (12 SOCIAL-BAP and 14 RIRE-BAP items, 9 of them reintegrated in this step) conforming to the two related dimensions without cross-loadings. As expected, the fit indices of this model (see Table 3) led us to revise it by discarding those malfunctioning items in each group according to their MIs. Afterwards, the resulting model Step.2.c.Model2 (11 SOCIAL-BAP and 11 RIRE-BAP items) presented adequate fit indicators in each sample (fit indicators in Table 3). We then aimed to perform the final item selection according to multigroup metric invariance on Step.2.c.Model2$_{\mathrm{MINV}}$. In this model, two items (SA3A3 and SC4C1i) with a considerable lack of invariance were discarded. After that, Step.2.c.Model3 (9 SOCIAL-BAP and 11 RIRE-BAP items) was compared with Step.2.c.Model3$_{\mathrm{MINV}}$ (multigroup metric invariance on the Step2.c.Model3), resulting in a lack of invariance between the Spanish samples and the UK sample (fit indices and models contrasts are presented in Table 3). Because the current content of the model fitted well in each sample separately and, as stated in the introduction section, some cultural differences were expected in the expression of the BAP, we decided to test a model with partial invariance regarding those items showing local misfit (Step.2.c.Model3$_{\mathrm{PMINV}}$, SB4B1, RC2A3, RC2B5, and RC2C3 factor loading were freed for the UK sample). We thus achieved the final BAP-IT content (20 items, nine for SOCIAL-BAP and 11 for RIRE-BAP) with metric invariance between the two different Spanish samples and with partial metric invariance between them and the English sample. Fit indices and model contrasts are presented in Table 3. The codes of the final BAP-IT items, the standardised and unstandardised factor loadings, and the intercepts of the model are displayed in Table 4.

The following step was to study if it was possible to assume structural invariance (i.e. an equal correlation between factors among samples). We first tried to seek structural invariance among the three groups. MI boundaries suggested differences in the correlation of the latent variables in the EN group. The model was respecified and structural invariance was fixed between the ES1 and ES2 groups but not with the EN group (Step.2.c.Model4). As the difference in the Chi-Square between this model and Step.2.c.Model$_{PMINV}$ was not significant (p > .05), Step.2.c.Model4 (invariant correlation between latent variables between ES1 and ES2) was accepted (fit indicators for Step.2.c.Model4 [structural invariance] and contrasts between these models are presented in Table 3). The standardised correlations among latent variables were .52 (p < .001) for ES1, .61 (p < .001) for ES2 group and .70 (p < .001) for EN. Figure 1 depicts the final BAP-IT model (Step.2.c.Model4) with standardised and unstandardised factor loadings. The content of the items in the final model is displayed in Table 5.

Figure 1

Standardized Factor Loadings [Non-Standardized] and Correlations of the Final Step.2.C.Model4 Model With Two Domains (SOCIAL-BAP And RIRE-BAP) in the Three Samples (ES1, ES2, And EN)



Note. Latent variable colours determine, based on their relationship, the colour of the manifest variables (items). The relation of the non-invariant items with the factor in the EN sample are coloured with red arrows. All factor loadings are statistically significant.

Additionally, to confirm the stability of the obtained results, we proceed to replicate, twice, the Step.2.c.Model4 and its contrast (Step.2.c.Model$_{PMINV}$) respectively, in the two randomly assigned half-samples of the original groups (each data set contained 981 and 983 complete cases each from ES1 [n_1 = 485, n_2 = 485], ES2 [n_1 = 230; n_2 = 231], and EN [n_1 = 266; n_2 = 267]). The results allowed us to confirm the stability of the BAP-IT measurement model in both languages (fit indices and contrasts between these models are presented in Table 3).

Table 5

BAP-IT Instructions and Items of the Spanish and English Versions

BAP-IT-ES	**BAP-IT-EN**
Para cada pregunta, marque con una (X) el valor que mejor describa con qué frecuencia esa frase se aplica a usted según los valores de la siguiente lista:	For each question, mark with a (X) the answer that best describes how often that statement applies to you on a scale of:
1. Nunca / 2. Casi Nunca / 3. Ocasionalmente / 4. A menudo / 5. Casi Siempre / 6. Siempre	1. Never / 2.Almost Never / 3. Occasionally / 4. Often / 5. Almost Always / 6. Always
SA1A7. En situaciones sociales, **no** se me ocurren temas de conversación.	SA1A7. In social situations, I **cannot** think of any topics of conversation.
SA2A2. En los descansos de una actividad grupal (ej., trabajo, clase, etc.) prefiero quedarme haciendo mis cosas (ej., adelantar trabajo, leer, etc.) antes que hablar con el resto de mis compañeros/as.	SA2A2. During breaks in group activities (e.g., work, class, etc.), I prefer to do my own thing (e.g., get a head-start on work, read, etc.) to speaking to the other people in the group.
SA3A7. Cuando una persona desconocida me habla, me limito a responder de manera breve a sus preguntas.	SA3A7. When a stranger speaks to me, I only give brief answers to their questions.
SA3A8. Cuando voy en un transporte público y alguien se sienta a mi lado, finjo estar ocupado/a (ej., cojo el móvil) para que **no** me hablen.	SA3A8. When I travel on public transportation and someone sits beside me, I pretend to be busy (e.g., I take out my mobile) so they will **not** speak to me.
SA3A11. Las conversaciones con personas desconocidas **no** me interesan.	SA3A11. Conversations with people I do **not** know do **not** interest me.
SB4B1. Cuando estoy hablando con alguien, **no** tengo claro hacia dónde mirar.	**SB4B1**. When I am talking to someone, I am **not** sure where I should look.
SRS60. Los demás piensan que soy emocionalmente distante y que **no** muestro mis sentimientos.	SRS60. Other people think I am emotionally distant and do **not** show my feelings.
SC3B5. Me molesta que mis amigos/as inviten a personas que **no** conozco a nuestras quedadas.	SC3B5. I get annoyed when my friends invite people I do **not** know to our gatherings.
SC4D1. Cuando acudo a un evento social donde **no** conozco a todo el mundo, me pone nervioso/a **no** saber qué decir ni qué hacer.	SC4D1. When I attend a social event where I do **not** know everyone, I get worried about **not** knowing what to say or do.
RA2A10. Mientras hablo, alineo, ordeno o clasifico las cosas que tengo al alcance de mi mano de forma casi involuntaria.	RA2A10. While I am speaking, I align, organise or classify objects within reach almost involuntarily.
RA3A3. Cuando estoy nervioso/a y a solas, digo en voz alta lo primero que me viene a la cabeza.	RA3A3. When I am nervous (anxious) and alone, I say out loud the first thing that comes to mind.
RA3A6. Cuando escucho una frase que me gusta, la repito literalmente en otras ocasiones, aunque **no** se ajuste al contexto	RA3A6. When I hear a phrase that I like, I repeat it verbatim on other occasions even though it does **not** fit the context.
RB3B2. Antes de salir de casa, me preparo de forma concreta para evitar salir con la sensación de haber olvidado algo.	RB3B2. Before leaving home, I have a specific way of preparing to avoid leaving with the feeling I have forgotten something.

BAP-IT-ES	BAP-IT-EN
RB4A2. Independientemente de la situación, saludo de la manera en que me siento más cómodo/a.	RB4A2. Regardless of the context, I use my own ways of greeting (e.g., sayings, set phrases) instead of conventional greetings.
RC2A3. Me encantan los objetos y juegos mecánicos donde el movimiento de unas partes afecta al de otras (ej., trenes, coches, objetos en equilibrio, bolas que chocan, etc.).	**RC2A3.** I love objects and mechanical games in which the movement of certain parts affects others (e.g., trains, cars, balance toys, pendulum balls, etc.).
RC2B5. Me fijo en todos los números con los que me encuentro (ej., las matrículas de los coches, las fechas, los números de carnés, etc.).	**RC2B5.** I pay close attention to all the numbers I come across (e.g., number plates on cars, dates, identity card numbers, etc.).
RC2C3. Me resisto a desprenderme de objetos que otras personas **no** valoran.	**RC2C3.** I have a hard time getting rid of objects which other people usually attach **no** importance to.
RD1A5. Me tomo la comida sin importarme si está muy caliente.	RD1A5. I eat food without caring whether it is very hot.
RD3A7. Me detengo a oler cada uno de los alimentos que me voy a comer.	RD3A7. I take time to smell each one of the foods I am going to eat.
RD3A14. Tocar determinados objetos de forma repetida me relaja.	RD3A14. Touching specific objects repeatedly relaxes me.

Note. Boldface indicates non-invariant items between ES groups and EN one.
Negative marks (e.g., not) in items are highlighted in the test

Step 3: SOCIAL-BAP and RIRE-BAP score reliability

Omega coefficients were estimated to check score reliability in SOCIAL-BAP and RIRE-BAP (McDonald, 1999; Raykov, 1997). Omega estimation for each dimension accounted for all of the common variance in their items (i.e. the variance was due to both specific and global factors). Adequate omega values should be higher than .80 (Raykov & Marcoulides, 2011). As the ES1 and ES2 groups were demonstrated to be fully invariant, only one estimation for each dimension was conducted. Omega estimations [95% CI] were .88 [.87, .90] for SOCIAL-BAP and .87 [.85, .89] for RIRE-BAP in ES1 and ES2, and .90 [.89, .91] and .90 [.88, .92] respectively in EN.

Validity Evidence Based on the relationship of BAP-IT dimensions with Related Subscales

Step 1: BAP-IT Relations with the BAPQ Subscales (ES1, ES2, and EN samples)

The resulting measurement model (Step.2.c.Model4) was included in a multi-group structural model (ES1, ES2 and EN). The Aloof, Pragmatic Language and Rigid subscales (BAPQ) were also included in the model to study their connection to latent variables in the SOCIAL-BAP and RIRE-BAP. The Cronbach's α coefficient for the

Aloof subscales was .88 (ES1), .90 (ES2) and .91 (EN); for Pragmatic Language, it was .85 (ES1), .71 (ES2) and .84 (EN); and for Rigid, it was .87 (ES1), .88 (ES2) and .88 (EN).[15] Factor variances were fixed to one, factor means to zero and factor covariance was freely estimated. Correlations among the SOCIAL-BAP, RIRE-BAP, and the Aloof, Rigid, and Pragmatic Language subscale scores of the BAPQ were fixed equally among groups. MIs suggested differences in correlations of the latent variables among groups. We started with a fully constrained model in which all correlations among latent variables and groups were fixed to equality. The model was respecified several times (MacCallum, 1986) computing the Satorra-Bentler Scaled Chi-square every time a parameter was freed (2001) to calculate chi-square difference testing using the MLR estimator. After several respecifications, structural invariance was achieved in some of the correlations of the latent variables among groups.

Fit indicators for the model were: χ^2 (711) = 1968.285, $p < .001$, the scaling correlation factor for MLR was 1.22 and the chi-square contribution was 910.794 from ES1, 478.244 from ES2 and 579.247 from EN. RMSEA was .05 [.05, .06] and CFI = .91 and TLI = .90. This model was compared with one in which all of the correlations among the latent variables of the BAP-IT and BAPQ were freely estimated across groups. Chi-Square difference testing according to the Satorra-Bentler Scaled Chi-Square were: CD = 1.22, TRd = 12.30, Δdf = 10.00, $p = .27$. The SOCIAL-BAP and RIRE-BAP correlation and SOCIAL-BAP and Aloof (BAPQ) were equal between ES1 and ES2 but differed from those for EN; the SOCIAL-BAP and Pragmatic Language (BAPQ) correlation was equal between ES1 and EN; SOCIAL-BAP and Rigid (BAPQ) was equal in ES2 and EN. The RIRE-BAP and Rigid (BAPQ) correlation was equal in ES1 and EN, but not in ES2. Aloof (BAPQ) and Pragmatic Language (BAPQ) was equal in ES2 and EN; Aloof (BAPQ) and Rigid (BAPQ) and also Pragmatic Language (BAPQ) and Rigid (BAPQ) were equal between ES1 and EN. The correlations between the BAP-IT domains and the BAPQ subscales in each sample are presented in Table 6. The SOCIAL-BAP was moderate-to-high and positively related with all BAPQ subscales in the three groups, while the RIRE-BAP was low-

[15] Direct scores instead of factor ones were used in this study. This decision was primarily motivated by previous studies that have reported inadequate fit values for the three-factor solution of the BAPQ either in its English (Broderick et al., 2015) or Spanish (Godoy-Giménez et al., 2018) versions.

to-moderate and positively related with all BAPQ subscales. The highest relationship was that between SOCIAL-BAP and Aloof. The correlations between the BAPQ subscales were not stable across samples. The Aloof-Pragmatic Language correlation (BAPQ) was the same for ES2 and EN but differed for ES1, the Aloof-Rigid correlation (BAPQ) was stable between ES1 and EN and the Rigid-Pragmatic Language (BAPQ) correlation was stable between ES1 and EN. No correlation between the BAPQ subscales was stable between both Spanish groups.

Table 6

BAP-IT Subscale Correlations with BAP Ones in the Three Samples

		SOCIAL-BAP	RIRE-BAP	Aloof	Pragmatic Language
	ES1				
BAP-IT	SOCIAL-BAP				
	RIRE-BAP	.34[.03]			
BAPQ	Aloof	.72[.03]	.23[.02]		
	Pragmatic Language	.55[.03]	.35[.03]	.45[.02]	
	Rigid	.50[.03]	.32[.02]	.46[.02]	.37[.02]
	ES2				
BAP-IT	SOCIAL-BAP				
	RIRE-BAP	.34[.03]			
BAPQ	Aloof	.72[.03]	.32[.03]		
	Pragmatic Language	.37[.02]	.27[.02]	.37[.02]	
	Rigid	.52[.03]	.31[.03]	.60[.03]	.30[.02]
	EN				
BAP-IT	SOCIAL-BAP				
	RIRE-BAP	.58[.04]			
BAPQ	Aloof	.56[.03]	.23[.02]		
	Pragmatic Language	.55[.03]	.53[.03]	.37[.02]	
	Rigid	.52[.05]	.42[.03]	.46[.02]	.37[.02]

Step 2: Further Validity Evidence Based on BAP-IT Relations with Other Tests (BAPQ, RCS, and SRS-II; ES2 sample)

The SOCIAL-BAP and RIRE-BAP factor scores and their determinacies of the Step.2.c.Model4 for the ES2 group were estimated in MPlus v7.11 (Muthén & Muthén, 2012). The factor score determinacies were SOCIAL-BAP = .93 and RIRE-BAP= .90, indicating an acceptable level (values should be at least .80 and higher than .90 for some situations like factor scores used as dependent variables; Gorsuch, 1983).

A partial correlation network was displayed with the subscales of the BAP-IT, BAPQ, RCS and SRS-II using the R-package *bootnet* (Epskamp et al., 2017) applying the gLasso algorithm (Friedman et al., 2014) and EBIC (Chen & Chen, 2009) for the ES2 group. The edges were refitted without LASSO regularisation (Epskamp & Fried, 2018, p.12). We calculated the percentage of explained variance of each node in the network (Haslbeck & Fried, 2017; see Table 7). The use of this approach to explore the empirical partial correlations between variables allowed us to compare the predictive capability between the BAP-IT and the BAPQ subscale scores (i.e. the one presenting higher correlations with the same related variance will capture this part of the shared variance). Nevertheless, the exploration of both the partial and the zero-order correlations is indispensable.

Table 7

Total Correlations (Above the Diagonal) and Partial Correlations (Below the Diagonal) Among the Subscales of the Tests Represented in the Figure 2

		SOCIAL	RIRE	ALOOF	P.LAN	RIGID	R.SEEK	E.REAC	S.T.FO	C.RIGID	AWARE	COGNI	COMMUN	MOTIV	RRB
BAP-IT	**SOCIAL**	1.00	.68	.73	.60	.62	.54	.54	.58	.16	.47	.60	.71	.75	.60
	RIRE	.29	1.00	.44	.56	.52	.40	.42	.45	.05	.43	.59	.58	.52	.64
BAPQ	**ALOOF**	.29	_	1.00	.48	.60	.58	.43	.50	.16	.48	.55	.65	.73	.46
	P.LANG	.13	.12	_	1.00	.46	.36	.32	.37	.07	.57	.62	.67	.53	.58
	RIGID	.08	.06	.07	_	1.00	.68	.63	.71	.31	.42	.59	.56	.54	.54
RCS	**R.SEEK**	_	_	.17	_	.27	1.00	.61	.68	.28	.37	.50	.49	.55	.43
	E.REAC	.10	.05	_	_	.08	.10	1.00	.78	.35	.25	.47	.40	.51	.40
	S.T.FO	-.01	_	_	_	.21	.21	.52	1.00	.39	.36	.53	.50	.59	.48
	C.RIGID	_	-.15	_	_	.09	_	.08	.17	1.00	.08	.12	.13	.15	.15
SRS-II	**AWARE**	_	_	.03	.14	_	_	-.12	_	_	1.00	.66	71	.55	.62
	COGNI	_	.05	_	.09	.12	0.02	_	.05	_	.18	1.00	.80	.67	.71
	COMMUN	.06	_	.02	.16	_	_	_	_	_	.25	.33	1.00	.78	.76
	MOTIV	.22	_	.29	_	_	0.01	_	.10	_	_	-.01	.34	1.00	.64
	RRB	_	.23	_	.02	.04	_	_	.01	_	.11	.15	.30	.00	1.00

Note. BAP-IT subscales: SOCIAL (SOCIAL-BAP), RIRE (RIRE-BAP); BAPQ subscales: ALOOF, P.LAN (Pragmatic Language), RIGID; RCS subscales: R.SEEK (Routine Seeking), E.REAC (Emotional Reaction), S.T.FO (Short-term Focus), C.RIGID (Cognitive Rigidity); SRS-II: AWARE (Social Awareness), COGNI (Social Cognition), COMMUN (Social Communication), MOTIV (Social Motivation), RRB (Restricted Repetitive Behaviours).

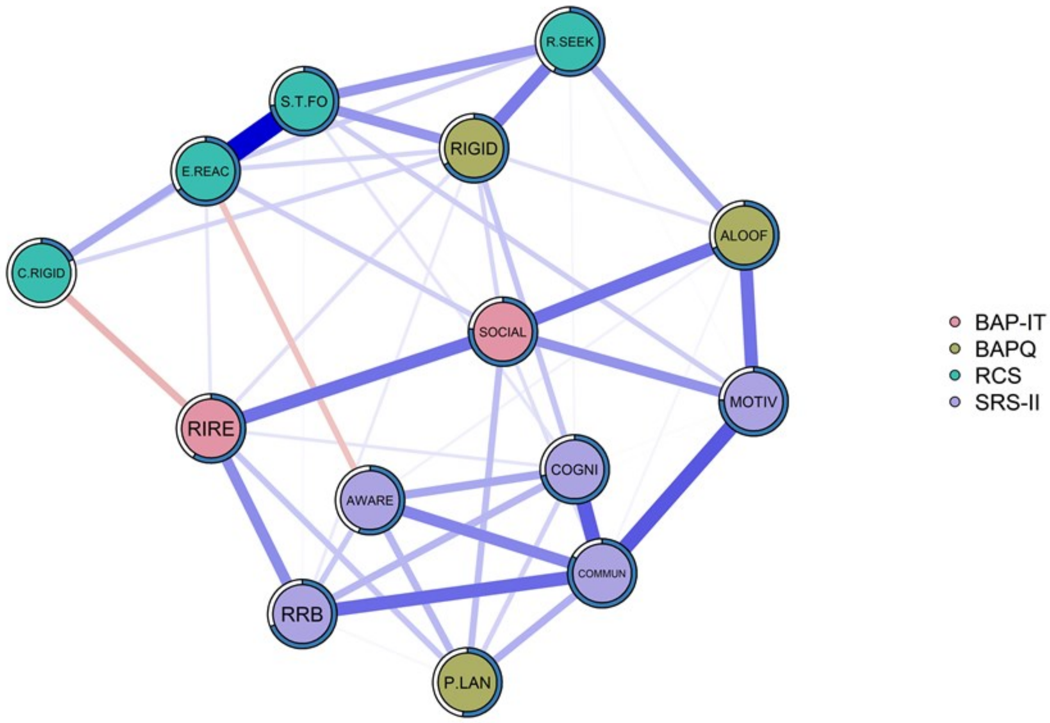
Table 7 shows the zero-order correlations in conjunction with the partial ones. Regarding zero-order correlations, the strongest connections of the SOCIAL-BAP were with Social Motivation (.75; SRS-II), Aloof (.73; BAPQ) and Social Communication (.71; SRS-II). It was also correlated with the rest of the subscale scores moderately, except in the particular case of Cognitive Rigidity (RCS), for which the correlation with SOCIAL-BAP was low (.16; see Table 7). All of the mentioned correlations were positive. RIRE-BAP highest correlations were with RRB (.64; SRS-II), Social Cognition (.59; SRs-II) and Social Communication (.58; SRS-II), followed by moderate and positive correlations with the rest of the subscales and a smaller correlation with Cognitive Rigidity (.05; RCS). Apart from its relationship with SOCIAL-BAP, the highest correlations of Aloof were with Social Motivation (.73; SRS-II) and Social Communication (.65; SRS-II). The highest correlations of Pragmatic Language were with Social Communication (.67; SRS-II) and Social Cognition (.62; SRS-II). Finally, the highest correlations of Rigid (BAPQ) were with Short-term Focus (.71; RCS) and Routine Seeking (.68; RCS). The SRS-II subscales and the RCS ones were moderate-to-highly correlated with the rest of the subscales of the test (see Table 7).

The resulting network is presented in Figure 2. Empirical partial correlations allowed us to observe which variable better explains other variables' variance (predictive capability). According to this, the highest connection of SOCIAL-BAP was with Aloof (BAPQ), followed by RIRE-BAP, Social Motivation (SRS-II), Pragmatic Language (BAPQ), Emotional Reaction (RCS), Rigid (BAPQ) and Social Communication (SRS-II); the strongest connection of RIRE-BAP was with SOCIAL-BAP, followed by RBB (SRS-II), Pragmatic Language (BAPQ), Rigid (BAPQ), Emotional Reaction (RCS) and Social Cognition (SRS-II). On the other hand, Aloof (BAPQ) presented connections to SOCIAL-BAP, Social Motivation (SRS-II), Routine Seeking (RCS) and Rigid (BAPQ). Pragmatic Language (BAPQ) showed relationships with Social Communication (SRS-II), Social Awareness (SRS-II), SOCIAL-BAP, RIRE-BAP and Social Cognition (SRS-II). For its part, Rigid (BAPQ) was related to Routine Seeking (RCS), Short-term Focus (RCS), Social Cognition (SRS-II), Cognitive Rigidity

(RCS),[16] SOCIAL-BAP, RIRE-BAP and Aloof (BAPQ). The rest of the SRS subscales referring to social deficits (i.e. Social Cognition and Social Awareness) were connected indirectly to SOCIAL-BAP via the Pragmatic Language subscale (BAPQ). Partial correlations can be seen in Table 7.

Figure 2

Empirical Network Model of the BAP-IT, BAPQ, SRS-II, and RCS Subscales



Note. Nodes represent test subscales. Nodes are connected by edges varying in thickness (thicker ages inform about greater partial correlations between subscales) and colour (blue/red edges denote positive/negative correlations respectively). The blue pie chart surrounding each node informs about its percentage of explained variance (R^2; more blue denotes more predicted variance). The colour of each node tells the test which the subscale belongs to. representing the subscales of each test are coloured regarding their test as displayed in the legend. BAP-IT subscales: SOCIAL (SOCIAL-BAP), RIRE (RIRE-BAP); BAPQ subscales: ALOOF, P.LAN (Pragmatic Language), RIGID; RCS subscales: R.SEEK (Routine Seeking), E.REAC (Emotional Reaction), S.T.FO (Short-term Focus), C.RIGID (Cognitive Rigidity); SRS-II: AWARE (Social Awareness), COGNI (Social Cognition), COMMUN (Social Communication), MOTIV (Social Motivation), RRB (Restricted Repetitive Behaviours).

[16] The zero-order correlation between RIRE-BAP and Cognitive Rigidity was .05. Thus, the negative partial correlation (−.15) could be due to a random variation close to the 0 value or be a consequence of an statistical effect (see Falk & Miller, 1992).

Discussion

The main aim of the present work was to develop a test for the two-dimensional operationalisation of the BAP proposed by Godoy-Giménez et al. (2021a) in two countries, Spain and the United Kingdom, based on the current ASD definition (DSM-5, APA, 2013; CID-11, WHO, 2019). The final BAP-IT is comprised 20 indicators distributed in two main subscales (SOCIAL-BAP and RIRE-BAP; SCI and RRB domains in Godoy-Giménez et al., 2021a).

The findings of the present study successfully report the BAP-IT as the only test created upon the basis of the two-dimensional structure of the BAP. Accordingly, evidence based on the internal structure of the test has been provided in two languages (English and Spanish). In addition to this, the BAP-IT is the only test that has followed a theory-driven approach and has tried to ensure that all of the central subdomains of the BAP (and not others) are represented. This solution has been shown to be fully metric-invariant between the Spanish groups and partially metric-invariant with the English group. Evidence of validity based on the relations of the BAP-IT dimensions with other variables has been also obtained. Finally, we would like to highlight that the BAP-IT is the first test with a two-dimensional structure in which items have been empirically chosen simultaneously in Spanish and English.

The BAP-IT thus also has a utility advantage, as it fills a gap in the measurement of the lower levels of autism in Spanish-speaking countries.

As the BAP-IT represents an attempt to solve previous concerns regarding the BAP psychometric assessment (i.e. adverse evidence regarding internal structure or contents of other tests which have been applied in the field; see Table 1), the findings are subsequently disclosed in greater detail. Regarding psychometric concerns, in the first place, we have provided evidence of validity based on the internal structure of the BAP and have further supported the dimensional operationalisation of the BAP (Godoy-Giménez et al., 2021a), which reconnected the phenotype with the disorder and reinserted it inside the autism continuum. Considering adverse evidence of the unstable internal structures of the former BAP tests, which is a recurrent topic in the psychometric assessment of psychological phenomena (Flake et al., 2017), here we opted to develop a test that would assure the stability of its structure in subsequent applications.

Specifically, from a candidate pool of 158, only 20 items were selected. One of the specific reasons for this substantial reduction is connected with the high presence of cross-loadings in several items (they were loading in both SOCIAL-BAP and RIRE-BAP). For instance, SB items shared variance with RIRE-BAP items, and the same could be applied to the RC items. We observed that the SB subdomain included deficits in nonverbal communicative behaviours (e.g. poorly integrated communication, abnormalities in eye contact and body language or problems understanding and using gestures) that were highly connected to echolalia and idiosyncratic speech from RIRE-BAP. Thus, although some of the items were built to represent SB (e.g. "I adjust my voice to emphasise what I am going to say next"), they could also have included aspects of this repetitive speech. Contrary to previous research suggesting that communication impairment was more closely related to social interaction deficits than to Restricted and Repetitive Behaviours (for a review, see Gernsbacher et al., 2016; see also Dworzynski et al., 2009; Lecavalier et al., 2006), our findings pointed to a strong connection between SB (non-verbal aspects of communication) and RA (repetitive and restricted speech). In this regard, other studies have lent a hand to ASD structure (APA, 2013) and suggested that echolalia is associated with sameness (Sterponi & Shankey, 2012). Even Kanner (1943) recognised that echolalia is characterised, overall, by rigidity and obsession and linked to asocial preoccupations (see also Rutter, 1978). Importantly, the observed link between SB and RA may be supported by studies that have described that echolalia and other repetitive speech may respond to interactional approaches highlighting the communicative value of many echoes (Sterponi & Shankey, 2012). Others have suggested that echolalia contributes to the formation and maintenance of social and emotional attachment and relationships (for a review, see Golysheva, 2019). The relations of BAP-IT with BAPQ subscales and the ES2 network (see below) might be also congruent with this reasoning. Validity evidence based on the BAP-IT relationships with other tests has shown that the Pragmatic Language subscale (BAPQ), which includes social communication (e.g. I am good at making small talk) and circumscribed patterns of communicating (e.g. I have been told that I talk too much about certain topics), is related to both SOCIAL-BAP and RIRE-BAP. Although the subdomains have demonstrated great relation even across dimensions (e.g. SB is related to RIRE-BAP subdomains as abovementioned), to obtain scores that can be easily interpreted, items should be unidimensional and assess only one domain at a time. Coherently to this psychometric premise and to assure structure stability, items that loaded in more than one dimension were eliminated.

Another possible cause for the large reduction in items is the great local dependency attributable to the similarity of the item content targeting each kind of behaviour/preference within each subdomain. At the heart of the BAP operationalisation (or any other construct) and the BAP-IT construction (or any other test), there is the need to capture the variance that is shared between different sorts of contents (behaviours and preferences) – that is, the similarities between behaviours that are only attributable to the construct and not to the shared content/wording of the items. After deleting items with cross-loadings, local dependency problems were disclosed, as some items started forming smaller groups sharing variance from a different source to the latent dimension of either SOCIAL-BAP or RIRE-BAP. This could have happened with several RC items, which also included aspects of social comparison (e.g. "When I lose or break something, I worry more than other people" or "I am sensitive to criticism").

Once we achieved a two-dimensional solution, a thorough inspection throughout the analysis revealed that, in SOCIAL-BAP, the items with the highest loadings were items from SC, but items from the three subdomains showed high loadings (many of them above .60). In the case of RIRE-BAP, the subdomains with the highest loadings were RA and RD, and all the subdomains included items with moderate-to-high loadings (above .40 most of them; see Appendix 5, Tables 6 and 7). After item deletion, the remaining items showed a unique common source of variance, being either SOCIAL-BAP or RIRE-BAP. The BAP-IT was also fully invariant between ES1 and ES2, and this also represented an improvement concerning prior tests, which have shown structure instability when applied (see Table 1).

Additionally, the cross-validation process has allowed us to confirm the stability of the BAP-IT measurement model and derived results in two randomised and independent subsamples in both languages. Concerning the invariant structure of the cross-cultural test, unlike precedent tests – which have failed in replicating internal structure in subsequent adaptations to other languages – the BAP-IT presented partial metric invariance between both languages. Vandenberg and Lance (2000) suggested the partial invariance strategy should not be employed when a large number of indicators are found noninvariant, and in our study, only 1 out of 9 items from SOCIAL-BAP and 3 of 11 from RIRE-BAP were not invariant. Complementarily, Byrne et al. (1989) suggested that group comparisons on latent means are possible even when some of the factor loadings vary. Getting to this satisfactory condition requires that the model comprehend numerous items with key contents of the latent dimension and at least one of these must be invariant. Importantly, this item must be different from the one that was fixed to 1.00 for identification purposes (Byrne et al., 1989). The fact that the internal structure of the test and its contents have been replicated in three groups and two different countries may suggest that the items of the dimensions measuring the latent dimension to the same extent. It is also worth mentioning that omega as estimation of the internal consistency of a scale suggests that a group of items comprised in the subscale and that were expected to measure a single latent construct, really do so. In the BAP-IT case, this suggests that key items from the different subdomains are producing similar results in the latent dimension (with values above .80 for either SOCIAL-BAP or RIRE-BAP and in both countries); this is especially important because the prior literature had reported diverse substructures in autistics domains (Georgiades et al., 2007; Lecavalier, 2005; Tadevosyan-Leyfer et al., 2003; Van Lang et al., 2006). Added to the fact that these similar results are stable across countries, this implies that the specific way an item relates to the dimension and the rest of the items involved in it are the same for both test versions, thus supporting the stability of the measurement structure even across countries. Another important strength concerns structural invariance, because we have obtained an invariant correlation in latent variables (SOCIAL-BAP and RIRE-BAP) between the groups ES1 and EN.

Finally, it should be noted that the three subdomains of the SOCIAL-BAP and four of the RIRE-BAP varying under the same dimension provides support not only for the internal structure of the BAP-IT but also to the operationalisation of the BAP and, by extension, the ASD criteria in the DSM-5 (APA, 2013). This is something extremely valuable and relevant. No prior BAP test has achieved this objective, because such tests are normally created following a data-driven approach and neglecting psychometric premises (Table 1). This links directly with the second problem we tried to amend: the lack of some central BAP contents.

Previously, Godoy-Giménez et al. (2021a) tried to collapse the items of the three most commonly applied tests for the assessment of the BAP and reshape them inside the new BAP operationalisation, determining that some central BAP behaviours were underrepresented in the studied measures. As the evidence argues in favour of the importance of those central behaviours in the ASD structure (APA, 2013; e.g. hyposensitivity and hypersensitivity autistic behaviours, Baum et al., 2015; unusual sensory interests are associated with more severe autistic behaviours, Zachor & Ben-Itzchak, 2013), the present work was primarily motivated to fill these gaps in the BAP measures. After careful work, all of the central BAP behaviours are presented in the BAP-IT counting on, at least, two items per subdomain, unlike previous tests (see Table 1). Thus, we could conclude that we have achieved one of our first motivations with this study.

The achievements of BAP-IT relative to the updated BAP definition representativeness and BAP structure could be the key to providing innumerable outcomes regarding the aetiology of autism condition, its spread in the population and its life development. Empirically – and linked to this – it opens the door to invaluable experimental situations in which varied samples can be reunited and compared (e.g. different ages, genders, socioeconomic status, countries, languages). Practically, it enables the testing of new assessment and intervention protocols. Of relevance, as the stability of the BAP-IT structure and its contents has also reported invariance, all of these advantages can now be fostered internationally, and this further enriches the field of research, because future studies could explore BAP expressions in two different countries and see whether people conceptualise the disorder similarly (Brown, 2015; Meredith, 1993). Additionally, after reaching invariance, the BAP-IT has opened the door to comparing the ratings of two different populations, as it is supposed that both populations are answering the items in identical ways (Chakraborty, 2017).

Once the final form of the BAP-IT was obtained, we focused on gathering evidence of validity based on dimensions relations to other variables. Regarding the first source of validity (BAP-IT relations with BAPQ subscales), the highest relationship was that between SOCIAL-BAP and Aloof (BAPQ). This connection was expected, as both comprehend key BAP behaviours. In the same way, Pragmatic Language (BAPQ) has been related to both SOCIAL-BAP and RIRE-BAP, which is in line with the fact that its items represent not only the social aspects of communication but also the repetitive patterns of speech so characteristic of autism (APA, 2013; Godoy-Giménez et al., 2021b [manuscript submitted to a journal]). Unexpectedly, an issue that will be recaptured later is the fact that RIRE-BAP and Rigid (BAPQ) relationship was moderate and lower than the SOCIAL-BAP and Rigid one. It is also worth mentioning that the SOCIAL-BAP and Aloof (BAPQ) relation was the same between the ES1 and ES2 samples; the SOCIAL-BAP and Pragmatic Language (BAPQ) relation between ES1 and EN; and the SOCIAL-BAP and Rigid (BAPQ) relation between ES2 and EN. However, no correlation between the latent variables of BAP-IT and BAPQ was invariant among the three groups. The relations among BAPQ subscales also showed variance among groups. Although some correlations between BAPQ subscales among groups were stable (e.g. the Aloof-Rigid correlation has shown invariance between ES1 and EN), none of the correlations remained equal between both Spanish groups, ES1 and ES2. Mostly considering the achieved structural invariance in the SOCIAL-BAP and RIRE-BAP relation between ES1 and ES2, this might indicate a problem in the stability of the internal structure of the BAPQ, which has been previously reported (e.g. Broderick et al., 2015; Godoy-Giménez et al., 2018; Lin et al., 2021; Sharma et al., 2018; Stojković et al., 2018). Future research should clarify this issue.

The additional evidence of validity studied in the ES2 sample has attempted to further explore these relationships between the BAP-IT and BAPQ subscales and of those subscales with other tests. As in the previous analysis, SOCIAL-BAP was related to Aloof and Pragmatic Language (BAPQ) and RIRE-BAP to Rigid and Pragmatic Language (BAPQ). The zero-order correlations and network provided evidence supporting the notion that SOCIAL-BAP may be assessing mild social communication and relation behaviours prototypical of autism throughout the general population. Additionally, we obtained evidence indicating that variations in RIRE-BAP could be attributable to key aspects of the second autistic dimension besides the fixated routines, which only contemplate two of the four subdomains. In contrast, Rigid (BAPQ) could be targeting – due to its more limited content (Godoy-Giménez et al., 2021a) –sameness and daily routines more specifically (e.g. “I am comfortable with unexpected changes in plans”). These findings matched our hypotheses.

A thorough look at these connections disclosed, following Constantino and Gruber (2012), that Social Communication (SRS-II) assesses reciprocal communication in social situations. Its positive and strong connection with SOCIAL-BAP and Pragmatic Language (BAPQ), and milder relation with Aloof (BAPQ), were as expected. Equally, the Social Motivation (SRS-II) subscale score, which represents the strongest relationships with SOCIAL-BAP and Aloof, refers to the degree to which an individual is positively predisposed (motivated) to participate in social interactions. Its high and similar connections to SOCIAL-BAP and Aloof suggest that variance regarding the SA subdomain (deficits in social-emotional reciprocity) and SC (deficits in developing, maintaining and understanding relationships) of the BAP have been adequately captured.

Furthermore, the correlations of Social Awareness (SRS-II) and Social Cognition (SRS-II) with SOCIAL-BAP and RIRE-BAP were moderate and similar to those found with the BAPQ subscales. Constantino and Gruber (2012) declared that both subscales are clustered in the first of the two domains in which the five subscales could be grasped to correspond to the DSM-5 diagnostic criteria for ASD. However, our findings suggested that Social Awareness and Social Cognition, but also Social Motivation and Social Communication, are related to both ASD dimensions. To further clarify the inherent nature of these subscales and their relations to either of the two ASD domains, more evidence is required. One possible solution may be to compare them with a different test that also measures interpretations of social situations such as the Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004).

Finally, RRB (SRS-II) addresses stereotyped and circumscribed interests and behaviours. The operationalisation of this subscale and the included content (e.g. "I have sensory interests that others find unusual [for example, smelling or looking at things in a special way]") adequately paralleled the second domain of ASD in the DSM-5 (APA, 2013). Correlations have uncovered a moderate-to-high relation with RIRE-BAP and a lower relationship with Rigid (this last even lower than the correlation of this RRB subscale with SOCIAL-BAP or Pragmatic Language [BAPQ]). In sum, this could represent an endorsement for the reconfiguration of the BAP structure carried out by Godoy-Giménez et al. (2021a).

Nevertheless, we also included another test that specifically targets restricted behaviours closely related to those comprehended in the RIRE-BAP. The RCS aims to assess *"dispositional resistance to change"* (Oreg, 2003), which refers to the inherent predisposition of some individuals to struggle with changes. Paying attention to zero-order correlations, the RIRE-BAP was related to all of the RCS subscales, having the lowest connection with Cognitive Rigidity (RCS) and the highest with Short-term Focus; a similar pattern was observed in the Rigid subscale [BAPQ]. While its relation to RRB may suggest that the RIRE-BAP correctly represents central behaviours of the RA (Stereotyped motor movements, use of objects, and speech; see Footnote 5) and RD (Hyper- or Hypo-reactivity to sensory input or unusual interests in sensory aspects) subdomains, its connections to all of the RCS subscales are appropriately targeting the RB (Insistence on sameness and inflexible adherence to routines) and RC (Highly restricted interests that are abnormal in intensity or focus) as well (RCS subscales seem to be related to RB and RC).

The case of the Cognitive Rigidity subscale (RCS) appears particularly relevant. In our view, additional evidence of validity based on the content of this subscale and its relations to other variables should be considered, bearing in mind that it is supposed to capture those autistic behaviours involving stubbornness and opposition to alternative ideas, perspectives, and methods (Oreg, 2003). This subscale has one of its strongest connections to Rigid (BAPQ; zero-order correlation was .31) but has shown very low connection to RIRE-BAP, low-to-moderate relations with the rest of the subscales of the RCS and all the variables included in the analysis. An alternative approach would also be to study the relations of RIRE-BAP with comorbid autistic behaviours. There is accumulative evidence suggesting that Obsessive Compulsive Disorder behaviours may be similar to the restricted and repetitive behaviours that are highly representative of and core to autism (e.g. Bolton et al., 1998; Micali et al., 2004; Wainer et al., 2011). The Obsessive Compulsive Inventory-Revised (Foa et al., 2002) targets such behaviours, and it could be interesting to study the relations of its subscales to those of BAP-IT.

The mentioned patterns of relationships together, with the empirical network findings, provide reasons for concluding that the SOCIAL-BAP may better capture the variance of those behaviours linked to the first domain of the DSM-5 (Impaired Social Communication and Social Relation; criterion A; APA, 2013) than Aloof (BAPQ). As substantive support for the found connections and in line with the work in which BAPQ was developed, it seems that Aloof ("a lack of interest in or enjoyment of social interaction"; Hurley et al., 2007, pp. 1681) might be more focused on this motivational engagement to social events (as referred in Social Motivation [SRS-II], which was the subscale with the strongest connection to Aloof). By contrast, the Pragmatic Language subscale (BAPQ) could rather be comprehended as the social impaired behaviours of autism than those originally intended ("pragmatic language problems refer to deficits in the social aspects of language, resulting in difficulties communicating effectively or in holding a fluid, reciprocal conversation"; Hurley et al., 2007, pp. 1681), as it shows its network connections to Social Communication and Social Awareness (SR-II) and its modulation in the indirect connections between those two subscales and SOCIAL-BAP.

As a second conclusion, the comparison of the RIRE-BAP and Rigid (BAPQ) subscales makes it possible to see how RIRE-BAP may better explain the variance of all of the central aspects of the second domain of autism (Restricted and Repetitive Patterns of Behaviours; criterion B; APA, 2013). Zero-order correlations supported this, as the relations of RRB with Rigid were moderate, and the same happened with the rest of the SRS-II subscales scores. This is in line with the empirical results found: repetitive behaviours seem to be key in RRB (SRS-II) and they have also been considered central behaviours of autism represented in the RA and RD subdomains. Of relevance here is the fact that the RRB subscale variance is better explained by RIRE-BAP. By contrast, Rigid (BAPQ) may be better at addressing the variance of the sameness-related behaviours than RIRE-BAP, as was shown by its partial correlations to all of the RCS subscales. This suggests that the Rigid subscale (BAPQ) may be better in predicting this variance. Nevertheless, the zero-order correlations let us conclude that this may not be adverse evidence of validity. The Rigid (BAPQ) definition clearly states the observed connections to Routine Seeking and Short-term Focus ([RCS]; see Hurley et al., 2007), which may suggest that this subscale is more focused on assessing the core rigid aspects of autism (sameness [Routine seeking] and extreme discomfort for inconveniences derived from change [Short-term Focus]) than the repetitive ones. In addition, Aloof (BAPQ) was the second subscale with the highest partial correlation with Routine Seeking (RCS), and from this, we can infer that the three BAPQ subscales are more closely linked to sameness (as supported by Godoy-Giménez et al., 2021a).

Finally and most importantly, we cannot omit that prior work has disclosed the problems associated with the BAPQ Pragmatic Language subscale (see Broderick et al., 2015; Godoy-Giménez et al., 2018), which drove the re-operationalisation of the BAP and the attempts to develop new assessment tools (see above or Godoy-Giménez et al., 2021a). Congruently, support for this also came from theory, as this third domain of the autism disorder (APA, 1994) was removed from the current ASD operationalisation and its contents have been coherently relocated inside the SB and RA subdomains of the first and the second domains of the ASD (criterion A and B; APA, 2013). As a strength of this paper, zero-order correlations of Pragmatic Language (BAPQ) with both SOCIAL-BAP and RIRE-BAP and with other subscales that refer to similar aspects (Social Communication and RRB [SRS-II]) agree with this. Likewise, the partial correlation between RIRE-BAP and Pragmatic Language (BAPQ) and between Pragmatic Language (BAPQ) and SOCIAL-BAP are almost the same, proposing that RIRE-BAP could be contemplating verbal stereotypies that Rigid (BAPQ) is not considering. Overall, it seems that the two-dimensional structure of the BAP proposed in the BAP-IT is not only more parsimonious but is also completely faithful to the current and worldwide accepted ASD operationalisation (APA, 2013). The BAP-IT better captures the variance attributable to most of the central behaviours of autism, as it was intended to do. This is consistent with the previous work by Godoy-Giménez et al. (2021a), in which a group of experts declared that the BAPQ subscales lacked RA and RD prototypical behaviours. In conclusion, all of the evidence provided here supports the dimensional structure of the BAP-IT and also the inferences made upon the BAP-IT scores.

As the BAP-IT aims to be a screening tool for application in the general population, we could conclude that we have provided empirical evidence supporting our purposes. Unlike other previous tests, one of the main strengths of our study is that we have attained a sample of key items (central autistic behaviours) representing the seven subdomains of the BAP with an invariant internal structure between samples of the same country and partially invariant between countries. The importance and utility of those central autistic components are clear, as has been discussed above. The BAP-IT thus represents a milestone in research on BAP and broadly in autism, because it enables the study of the two-dimensional BAP structure and allows for the comparison of both BAP and ASD constructs even between two different countries. The evidence of validity provided highlights the goodness of our test and the reliance on the inferences made on BAP-IT scores. The stability of the structure of the test across countries was called for and also matched the expected relations with tests that assess key autistic behaviours or closely related ones. Still, research on the BAP-IT has just started, because its validation will be an ongoing task (Flake et al., 2017). Alternative studies could contribute by setting the relational definition of the BAP (nomological net, Cronbach & Meehl, 1955; e.g. an aloof personality could be linked to shyness, Murphy et al., 2000) by providing further evidence of validity based on the relations of the BAP-IT subscales to other variables (e.g. Obsessive Compulsive Disorder behaviours could be linked to restricted and repetitive behaviours; Wainer et al., 2011). On the other hand, in line with the dimensional approach to autism and the new ASD severity specifiers (APA, 2013), future studies could explore how BAP characteristics behave similarly or differently across countries; item response models could be highly relevant in this regard (Bond & Fox, 2015; Hambleton, 2006). The application of the item response models could also help to achieve invariant structure for the BAP-IT across versions (e.g. by studying differential test or items functioning). This will open the door to international research on the BAP.

Acknowledgments. This work was supported by grants PID2019-110066GB-I00/MCIN/AEI/ 10.13039/501100011033 to AFE; and the Spanish Ministry of Education, Culture and Sport to MGG (FPU2015-015).

Author Contribution. All the authors took part in the Item Selection and Assignment According to their Content and Itemmetric Properties. MGG, PSJ, and AGR were responsible for data collection. PSJ, AFE, and FC were responsible for the design of the study. PSJ and MGG analysed the data and wrote the manuscript. All authors contributed to data interpretation and assisted with writing the manuscript. All the authors gave their consent for publishing this study.

Declarations

Conflicting interests. Authors have no known conflict of interest to disclose.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (DSM-IV®)*. American Psychiatric Pub.

American Educational Research Association. (2014). *Standards for educational and psychological testing*. American Educational Research Association American Psychological Association National Council on Measurement in Education.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV-TR®)*. American Psychiatric Pub.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Angleitner A., John O.P., Löhr FJ. (1986) It's *What* You Ask and *How* You Ask It: An Itemmetric Analysis of Personality Questionnaires. In: Angleitner A., Wiggins J.S. (eds) *Personality Assessment via Questionnaires.* Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-70751-3_5

Armstrong, T. S., Cohen, M. Z., Eriksen, L., & Cleeland, C. (2005). Content Validity of Self-Report Measurement Instruments: An Illustration From the Development of the Brain Tumor Module of the M.D. Anderson Symptom Inventory. *Oncology Nursing Forum, 32*(3), 669–676. https://doi.org/10.1188/05.onf.669-676

Asparouhov, T., & Muthén, B. (2010). Plausible values for latent variables using Mplus. *Unpublished. Accessed at: http://www. statmodel. com/download/Plausible. pdf.*

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of autism and developmental disorders*, *31*, 5-17. https://doi.org/10.1023/A:1005653411471

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders, 34*(2), 163-175. https://doi.org/10.1023/B:JADD.0000022607.19833.00

Baum, S. H., Stevenson, R. A., & Wallace, M. T. (2015). Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. *Progress in Neurobiology, 134,* 140–160. https://doi.org/10.1016/j.pneurobio.2015.09.007.

Beck, C. T., & Gable, R. K. (2001). Ensuring content validity: An illustration of the process. *Journal of Nursing Measurement*, *9*(2), 201-215. https://doi.org/10.1891/1061-3749.9.2.201

Bölte, S., Poustka, F., & Constantino, J. N. (2008). Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Research*, *1*(6), 354-363. https://doi.org/10.1002/aur.49

Bolton, P. F., Pickles, A., Murphy, M., & Rutter, M. (1998). Autism, affective and other psychiatric disorders: patterns of familial aggregation. *Psychological medicine, 28*(2), 385-395. https://doi.org/10.1017/S0033291797006004

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). L. Erlbaum.

Broderick, N., Wade, J. L., Meyer, J. P., Hull, M., & Reeve, R. E. (2015). Model invariance across genders of the broad autism phenotype questionnaire. *Journal of autism and developmental disorders, 45*(10). https://doi.org/3133-3147. 10.1007/s10803-015-2472-z

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychological bulletin, 105*(3), 456. https://doi.org/10.1037/0033-2909.105.3.456

Chakraborty, R. (2017). Configural, metric and scalar invariance measurement of academic delay of gratification scale. *Int. J. Human. Soc. Stud*, *3*(3). https://doi.org/10.20908/ijarsijhss.v3i3.7903

Chen, Z., & Chen, J. (2009). Tournament screening cum EBIC for feature selection with high-dimensional feature spaces. *Science in China Series A: Mathematics, 52(6)*, 1327-1341. https://doi.org/10.1007/s11425-009-0089-4

Chien, Y. L., Wu, Y. Y., Chiu, Y. N., Liu, S. K., Tsai, W. C., Lin, P. I., ... & Chien, W. H. (2011). Association study of the CNS patterning genes and autism in Han Chinese in Taiwan. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(6), 1512-1517. https://doi.org/10.1016/j.pnpbp.2011.04.010

Constantino, J. N., & Charman, T. (2016). Diagnosis of autism spectrum disorder: reconciling the syndrome, its diverse origins, and variation in expression. *The Lancet Neurology*, *15*(3), 279-291. https://doi.org/10.1016/S1474-4422(15)00151-9

Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale: SRS-2*. Torrance, CA: Western Psychological Services.

Constantino, J. N., & Gruber, C. P. (2017). *Social responsiveness scale Spanish: SRS-2*. Torrance, CA: Western Psychological Services.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Archives of general psychiatry*, *60*(5), 524-530. https://doi.org/10.1016/S0140-6736(18)31129-2

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, *52*(4), 281. https://doi.org/10.1037/h0040957

Delgado-Rico, E., Carretero-Dios, H., & Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology Espana*, *12*(3), 449-460. https://doi.org/10.5167/uzh-64551

Downing, S., & Haladyna, T. (2006). *Handbook of test development.* Lawrence Erlbaum.

Dworzynski, K., Happé, F., Bolton, P., & Ronald, A. (2009). Relationship between symptom domains in autism spectrum disorders: a population based twin study. *Journal of Autism and Developmental Disorders, 39*(8), 1197-1210. https://doi.org/10.1007/s10803-009-0736-1

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods, 50*, 195-212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological methods,* *23*(4), 617. https://doi.org/10.1037/met0000167

Falk, R., & Miller, N. (1992). *A Primer for Soft Modeling*. Akron, OH: The University of Akron Press.

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and*

Personality Science, 8(4), 370-378. https://doi.org/10.1177/1948550617693063

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics, 9*, 432–441. https://doi.org/10.1093/biostatis tics/kxm045

Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., Lee, A. B., ... & Buxbaum, J. D. (2014). Most genetic risk for autism resides with common variation. *Nature genetics*, *46*(8), 881-885. https://doi.org/10.1038/ng.3039

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., Roberts, W., ... & Mahoney, W. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(2), 188-196. https://doi.org/10.1097/01.chi.0000242236.90763.7f

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A newapproach for estimating the number of dimensions in psychological research. *PLoS One, 12*, e0174035. https://doi.org/10.1371/journal.pone.0174035

Godoy-Giménez, M., González-Rodríguez, A., Cañadas, F., Estévez, A. F., & Sayans-Jiménez, P. (2018). Psychometric Properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements. *Journal of autism and developmental disorders*, *48*(3), 770-783. https://doi.org/10.1007/s10803-017-3438-0

Godoy-Giménez, M., González-Rodríguez, A., Cañadas, F., Estévez, A. F., & Sayans-Jiménez, P. (2021a). Is it possible to assess the two-domain definition of the Broad Autism Phenotype using the available measurement tools? *Journal of autism and developmental disorders.* https://doi.org/10.1007/s10803-021-05158-7

Godoy-Giménez, M., González-Rodríguez, A., Estévez, A. F., Cañadas, F., & Sayans-Jiménez, P. (2021b). *Applying a Rasch Model to the BAPQ: Item Severity Analysis and Test Differential Functioning of the English and Spanish Versions* [Manuscript submitted for publication]. Research in developmental disabilities.

Golysheva, M. D. (2019). A Review on Echolalia in Childhood Autism. *Advances in Social Science, Education and Humanities Research, 333*, 200-204. https://doi.org/10.1007/s40489-015-0067-4

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in nursing & health*, *20*(3), 269-274. https://doi.org/10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.0.CO;2-G

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied measurement in education*, *15*(3), 309-333. https://doi.org/10.1207/S15324818AME1503_5

Hambleton, R. K. (2006). Good practices for identifying differential item functioning. *Medical care, 44*(11), S182-S188. https://doi.org/10.1097/01.mlr.0000245443.86671.c4

Hambleton, R. K., Swaminathan, H., Algina, J., & Coulson, D. B. (1978). Criterion-referenced testing and measurement: A review of technical issues and developments. *Review of Educational Research, 48*, 1-47. https://doi.org/10.3102/00346543048001001

Haslbeck, J. M. B., & Fried, E. I. (2017). How predictable are symptoms in psychopathological networks? A reanalysis of 18 published datasets. *Psychological medicine, 47*(16), 2767-2776. https://doi.org/10.1017/S003329171700125

Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological assessment*, *7*(3), 238. https://doi.org/10.1037/1040-3590.7.3.238

Hernandez, A., Hidalgo, M. D., Hambleton, R. K., & Gomez-Benito, J. (2020). International test commission guidelines for test adaptation: A criterion checklist. *Psicothema*, *32*(3), 390-398. https://doi.org/10.7334/psicothema2019.306

Hoekstra, R. A., Vinkhuyzen, A. A., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., ... & Van Der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). *Journal of autism and developmental disorders*, *41*(5), 589-596. http://dx.doi.org/10.1007/s10803-010-1073-0

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal, 6,* 1-55.

Hurley, R. S., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of autism and developmental disorders*, *37*(9), 1679-1690. http://dx.doi.org/10.1007/s10803-006-0299-3

Hurst, R. M., Mitchell, J. T., Kimbrel, N. A., Kwapil, T. K., & Nelson-Gray, R. O. (2007). Examination of the reliability and factor structure of the Autism Spectrum Quotient (AQ) in a non-clinical sample. *Personality and Individual Differences*, *43*(7), 1938-1949. https://doi.org/10.1016/j.paid.2007.06.012

IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp

International Test Commission. (2017). The ITC Guidelines for Translating and Adapting Tests (Second edition). [www.InTestCom.org]

Irwing, P., Booth, T., & Hughes, D. (2018). *The Wiley Handbook of Psycometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development.* John Wiley & Sons Ltd.

Jöreskog, K. G., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models.* Cambridge, MA.: Abt.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250.

Kloosterman, P. H., Keefer, K. V., Kelley, E. A., Summerfeldt, L. J., & Parker, J. D. (2011). Evaluation of the factor structure of the Autism-Spectrum Quotient. *Personality and Individual Differences*, *50*(2), 310-314. https://doi.org/10.1016/j.paid.2010.10.015

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2019). MVN: An R package for assessing multivariate normality. *The R Journal, 6*(2), 151-162.

Landry, O., & Chouinard, P. A. (2016). Why we should study the broader autism phenotype in typically developing populations. *Journal of Cognition and Development*, *17*(4), 584-595. https://doi.org/10.1080/15248372.2016.1200046

Lane, S., Raymond, M., & Haladyna, T. (2015). *Handbook of Test Development—2nd Edition.* Routledge.

Leamer, E. E. (1978). Regression selection strategies and revealed priors. *Journal of the American Statistical Association, 73(*363), 580-587. https://doi.org/10.1080/01621459.1978.10480058

Lecavalier, L. (2005). An evaluation of the Gilliam autism rating scale. *Journal of autism and developmental disorders, 35*(6), 795-805. https://doi.org/10.1007/s10803-005-0025-6

Lecavalier, L., Leone, S., & Wiltz, J. (2006). The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *Journal of intellectual disability research, 50*(3), 172-183. https://doi.org/10.1111/j.1365-2788.2005.00732.x

Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., ... & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature Reviews Disease Primers*, *6*(1), 1-23. https://doi.org/10.1038/s41572-019-0138-4

Lord, C., Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *The Lancet*, *392*(10146), 508-520. https://doi.org/10.1016/S0140-6736(18)31129-2

MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological bulletin*, *100*, 107. https://doi.org/10.1037/0033-2909.100.1.107

McDonald, R.P. (1999). *Test theory: A unified treatment*. Hillsdale: Erlbaum

Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*(3), 519-30. https://doi.org/10.1093/biomet/57.3.519

Meera, S. S., Girimaji, S. C., Seshadri, S. P., Philip, M., Shivashankar, N., Morgan, P., & Piven, J. (2015). Translation of the Broad Autism Phenotype Questionnaire to an Indian language: A description of the process. *Asian Journal of Psychiatry, 15*, 62–67. https://doi.org/10.1016/j.ajp.2015.04.013

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525-543. https://doi.org/10033-3123/93/1200-pa93500.75/0

Micali, N., Chakrabarti, S., & Fombonne, E. (2004). The broad autism phenotype: findings from an epidemiological survey. *Autism, 8*, 21-37. https://doi.org/10.1177/1362361304040636

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. International Journal of Psychological Research, 3(1), 111–121. dx.doi.org/10.21500/20112084.85

Morrison, K. E., Chambers, L. K., Faso, D. J., & Sasson, N. J. (2018). The content and function of interests in the broad autism phenotype. *Research in Autism Spectrum Disorders*, *49*, 25-33. https://doi.org/10.1016/j.rasd.2018.02.002

MPLUS (Version 7.11). [Computer Software]. Los Angeles, CA: Muthén & Muthén.

Murphy, M., Bolton, P. F., Pickles, A., Fombonne, E., Piven, J., & Rutter, M. (2000). Personality traits of the relatives of autistic probands. *Psychological Medicine, 30*(6), 1411-1424. http://dx.doi.org/10.1017/S0033291799002949

Muthén, B., & Asparouhov, T. (2011). Beyond multilevel regression modeling: Multilevel analysis in a general latent variable framework. In *Handbook of advanced multilevel analysis* (pp. 23-48). Routledge.

Muthén, L. K., & Muthén, B. O. (1998-2012). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Nayar, K., Sealock, J. M., Maltman, N., Bush, L., Cook, E. H., Davis, L. K., & Losh, M. (2021). Elevated polygenic burden for autism spectrum disorder is associated with the broad autism phenotype in mothers of individuals with autism spectrum disorder. *Biological psychiatry*, *89*(5), 476-485. https://doi.org/10.1016/j.biopsych.2020.08.029

Oreg, S. (2003). Resistance to change: Developing an individual differences measure. *Journal of applied psychology, 88*(4), 680. https://doi.org/10.1037/0021-9010.88.4.680

Osterlind, S. J. (2002). *Constructing Test Items: Multiple-Choice, Constructed-Response, Performance, and Other Formats*. Kluwer Academic Publishers. Retrieved from http://ebooks.kluweronline.com

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997a). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry, 154*(2), 185–190. http://dx.doi.org/10.1176/ajp.154.2.185

Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997b). Personality and language characteristics in parents from multiple-incidence autism families. American Journal of Medical Genetics, 74(4), 398-411.

https://doi.org/10.1002/(SICI)1096-8628(19970725)74:4<398::AID-AJMG11>3.0.CO;2-D

Pruitt, M. M., Rhoden, M., & Ekas, N. V. (2018). Relationship between the broad autism phenotype, social relationships and mental health for mothers of children with autism spectrum disorder. *Autism*, *22*(2), 171-180. https://doi.org/10.1177/1362361316669621

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate behavioral research, 32*(4), 329-353. https://doi.org/10.1207/s15327906mbr3204_2

Raykov, T., & Marcoulides, G. A. (2011). Procedures for estimating reliability. Introduction to psychometric theory. *New York: Routledge, Taylor & Francis Group,* 147-181.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happe, F., et al. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Archives of General Psychiatry*, *68*(11), 1113-1121. https://doi.org/10.1001/archgenpsychiatry.2011.119.

Rutter, M. (1978). Diagnosis and definition of childhood autism. *Journal of autism and childhood schizophrenia, 8*(2), 139-161. https://doi.org/10.1007/BF01537863

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514. https://doi.org/10.1007/BF02296192

Sasson, N. J., Lam, K. S. L., Childress, D., Parlier, M., Daniels, J. L., & Piven, J. (2013). The broad autism phenotype questionnaire: Prevalence and diagnostic classification. *Autism Research, 6*, 134–143. https://doi.org/10.1002/aur.1272

Shi, L.-J., Ou, J.-J., Gong, J.-B., Wang, S.-H., Zhou, Y.-Y., Zhu, F.-R., … Luo, X.-R. (2015). Broad autism phenotype features of Chinese parents with autistic children and their associations with severity of social impairment in probands. *BMC Psychiatry*, *15*(1). https://doi.org/10.1186/s12888-015-0568-9

Siu, W. K., Lam, C. W., Mak, C. M., Lau, E. T. K., Tang, M. H. Y., Tang, W. F., ... & Chan, K. Y. (2016). Diagnostic yield of array CGH in patients with autism spectrum disorder in Hong Kong. *Clinical and translational medicine*, *5*(1), 1-8. https://doi.org/10.1186/s40169-016-0098-1

Sörbom, D. (1989). Model modification. *Psychometrika, 54*(3), 371-384. https://doi.org/10.1007/BF02294623

Sterponi, L., & Shankey, J. (2014). Rethinking echolalia: Repetition as interactional resource in the communication of a child with autism. *Journal of child language, 41*(2), 275-304. https://doi.org/10.1017/S0305000912000682

Stewart, M. E., & Austin, E. J. (2009). The structure of the Autism-Spectrum Quotient (AQ): Evidence from a student sample in Scotland. *Personality and individual differences*, *47*(3), 224-228. https://doi.org/10.1016/j.paid.2009.03.004

Stewart, G. R., Charlton, R. A., & Wallace, G. L. (2018). Aging with elevated autistic traits: Cognitive functioning among older adults with the broad autism phenotype. *Research in Autism Spectrum Disorders*, *54*, 27-36. https://doi.org/10.1016/j.rasd.2018.06.009

Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B. R. I. A. N., Putnam, S. A. R. A., McGrath, L., ... & Folstein, S. E. (2003). A principal components analysis of the Autism Diagnostic Interview-Revised. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*(7), 864-872. https://doi.org/10.1097/01.CHI.0000046870.56865.90

Van Lang, N. D., Boomsma, A., Sytema, S., de Bildt, A. A., Kraijer, D. W., Ketelaars, C., & Minderaa, R. B. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *Journal of Child Psychology and Psychiatry, 47*, 37-44. https://doi.org/10.1111/j.1469-7610.2005.01434.x

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods, 3*, 4-70. https://doi.org/10.1177/109442810031002

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, *40*(5), 472-481. https://doi.org/10.1016/j.jrp.2005.03.003

Volkmar, F. R., Paul, R., Rogers, S. J., & Pelphrey, K. A. (Eds.). (2014). *Handbook of autism and pervasive developmental disorders, diagnosis, development, and brain mechanisms* (Vol. 1). John Wiley & Sons.

Wainer, A. L., Ingersoll, B. R., & Hopwood, C. J. (2011). The Structure and Nature of the Broader Autism Phenotype in a Non-Clinical Sample. *Journal of Psychopathology and Behavioral Assessment, 33*, 459–469. http://dx.doi.org/10.1007/s10862-011-9259-0

Waltz, C., & Bausell, R. B. (1983). *Nursing research: Design, statistics and computer analysis* (1st ed.). Philadelphia: F.A. Davis Co.

Wang, T., Guo, H., Xiong, B., Stessman, H. A., Wu, H., Coe, B. P., ... & Eichler, E. E. (2016). De novo genic mutations among a Chinese autism spectrum disorder cohort. *Nature communications*, *7*(1), 1-10. https://doi.org/10.1038/ncomms13316

Warrier, V., Chee, V., Smith, P., Chakrabarti, B., & Baron-Cohen, S. (2015). A comprehensive meta-analysis of common genetic variants in autism spectrum conditions. *Molecular autism*, *6*(1), 1-11. https://doi.org/10.1186/s13229-015-0041-0

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis.* MESA press.

World Health Organization. (2019). *International Statistical Classification of Diseases and Related Health Problems (CID-11®)*. World Health Organization.

Zachor, D. A., & Ben-Itzchak, E. (2013). The relationship between clinical presentation and unusual sensory interests in autism spectrum disorders: A preliminary investigation. *Journal of Autism and Developmental Disorders.* https://doi.org/10.1007/s10803-013-1867-y

BLOCK THREE:

CONCLUSIONS

Lo propio del saber no es ni ver ni demostrar, sino interpretar.

"Las palabras y las cosas"** (1966), **Michael Foucault

CHAPTER SEVEN. DISCUSSION OF THE OVERALL OUTCOMES AND CONCLUSIONS

Through the previous chapters of this dissertation, we have had the opportunity to review how different disciplines and theories have played a crucial role in such a complex task as it is the reconstruction of the BAP operationalisation and measurement. The aforementioned theoretical framework settled the foundations of this endeavour and the empirical chapters guided the scientific path for the construction of the dimensional operationalization of the BAP and its new test, the BAP-IT. The implications derived from this work are the main outcomes of this last block, which is divided into two chapters. In the first one, chapter seven, following the scientific approach, the overall results of my studies will be discussed and the main conclusions of the research will be drawn.

In the first place, I consider it necessary to briefly summarise the reasons that underlie the studies conducted through the present dissertation. The first part of this dissertation was motivated by the absence of a tool that targeted autism at BAP levels in adults in the general population in Spanish countries. Therefore, the timeline of my research starts with the adaptation of the BAPQ to Spanish.

The BAPQ-SP: The Spanish Version of the Original and Worldwide Applied Test of the BAP

In the first study, we presented the Spanish version of the BAPQ, the BAPQ-SP[17]. As already argued, the BAPQ was the original BAP test (aiming specifically for this construct) and it corresponds to the initial BAP operationalization which paralleled the autism disorder structure in the DSM-IV (APA, 1994). In the first place, our work with the BAPQ-SP and the BAPQ was focused on providing validity evidence based on the internal structure of the BAPQ-SP, and, subsequently, on the BAPQ subscales in relation to other variables (Study 1). In the second place, understanding that BAP and ASD are both expressions of autism at different levels of severity implies the expansion of the ASD continuum to capture BAP levels of expression. Yet, to follow this dimensional approach, we needed to test whether the items of the BAPQ-

[17] The translation of two items was corrected. An improved version of them can be consulted in Study 2.

SP (and the original BAPQ) which were supposed to represent autistic behaviours at BAP levels were scalable along the severity continuum of autism and that their severity adjusted to the BAP levels in the general population (Study 2). In the third place, in an attempt to explore BAP expressions cross-culturally, which was one of the main interests of the present dissertation, not only did we need to adapt a well-known test to the Spanish context, but we also needed to report test invariant functioning cross-culturally (Study 2).

Regarding the first objective of the present dissertation, providing validity evidence based on the internal structure of the BAPQ, it would have supported the BAP operationalization in this test. Furthermore, scores derived from its Spanish version would have represented the same construct as the original BAPQ, which would have enabled the comparability of the inferences made on test versions (AERA et al., 2014). However, once adapted, measurement problems started to arise as previously disclosed (i.e., Broderick et al., 2015; Sasson et al., 2013a). Analysis of the internal structure of the BAPQ (through factor analytic approaches; Study 1) revealed those suspected problems as a severe misfit in the three-factor model of the BAPQ, which was mainly caused by the Pragmatic Language subscale. Evidence based on BAPQ subscales relationships with other variables also contributed to this because some relationships of Aloof and Rigid were lower than expected. Furthermore, although the subscales of the Pragmatic Awareness Questionnaire (PAQ; Rodríguez-Muñoz, 2012) had been expected to be only related to the Pragmatic Language subscale, they were also connected to Aloof. This suggested that the Aloof subscale could also be including content related to non-verbal communicative behaviours contrary to what was expected according to the autism disorder definition in the DSM-IV (APA, 1994). Altogether, those results matched with misfit expectations and were in line with the disappearance of the Pragmatic Language domain in the ASD structure in the DSM-V. Relatedly, the observed connections of Pragmatic Language to both Aloof and Rigid subscales were also in line with this ASD structure. Thus, instead of collapsing Aloof and Pragmatic Language subscales into one (Sasson et al., 2013c), Pragmatic Language subscale content may be better relocated inside each of the two ASD domains.

The study of the ASD-BAP dimensional approach came from the item response theory (IRT) in Study 2. Importantly, results from this study have provided, again, unfavourable evidence regarding the Pragmatic Language subscale and, this time, in the BAPQ original version as well. In a similar fashion, lack of unidimensionality was observed in the Rigid subscale in the BAPQ-EN. Unsurprisingly, this is also congruent with most literature in the field (e.g., Lin et al., 2021; Stojković et al., 2018) and only reinforces accumulative unfavourable evidence based on the internal structure of the BAPQ. Aside from that, the unidimensionality of both versions of the Aloof subscale enabled us to propose two independent severity hierarchies and do the same for the Rigid subscale in the Spanish version. Inside each hierarchy (either Aloof or Rigid), items and persons were scaled in terms of the severity of behaviours they were expressing. Besides, both subscales (Aloof and Rigid) could statistically discern between high and low BAP levels. Yet, one important outcome is that order matched with the severity specifiers of domains A (Deficits in social communication and social interaction) and B (Patterns of restricted and repetitive behaviours and interests) of the ASD in the DSM-V (see Table 2, Chapter One, Block One; APA, 2013). For example, in Study 2, it is explained that a mild introverted behaviour could be to not enjoy being in social situations (item 9[i]: "*I enjoy being in social situations*"), which does not imply that the person avoids spending time with few close friends or relatives. On the contrary, expressing an instrumental preference for the acts of socializing, in general (i.e., also involves interactions with relatives, close friends, and partners) could be taken as a more severe indicator (item 5: "*I would rather talk to people to get information than to socialize*"). In the case of Rigid BAPQ-SP, the resulting order of the Rigid BAPQ-SP hierarchy could also open the door to insert behaviours at BAP levels of severity inside the autism continuum (De Groot & Van Strien, 2017; Hoekstra et al., 2008). In this sense, stubbornness, persisting routines and a strong need for sameness (i.e., Item 26: *"People get frustrated by my unwillingness to bend"*) could be closer to clinical levels of severity, mostly when inflexible behaviours cause significant interference with individual functioning in one or more contexts (Item 15[i]: *"I am flexible about how things should be done"*; level 1), it comprises difficulty coping with change (Item 8: *"I have to warm myself up to the idea of visiting an unfamiliar place"*; level 2), and those behaviours marking interfering in all the spheres (level 3; ASD severity specifiers; APA, 2013).

On the other hand, we were aware that the BAPQ was developed in a very specific and rather small sample (i.e., parents of children diagnosed with autism disorder; Hurley et al., 2007) and, later, applied to further samples such as people from the general population. For that reason, we needed to test whether the severity expressed by the items was matching the severity expressed in the targeted population. Following the dimensional hypothesis of autism (see Chapter One, Block One), parents and first-degree relatives of children diagnosed with ASD should present more intense autistic-like behaviours than the general population. As expected, the average severity of the items for both Aloof subscales in both test versions and Rigid BAPQ-SP was higher than the average severity level of their community samples. This suggested that subtler BAP indicators should be added to the test to enhance the measurement accuracy at lower BAP levels.

In the third place, bearing the international reportage of the BAP expressions and the authors' interest in understanding how autism behaves cross-culturally, which autistic expressions may result similar, and which do not, we considered that the Spanish version of the BAPQ should be thoroughly validated. The invariance between both versions of the same test is gauged by whether scores of the test subscales represent the same construct in the original and the adapted version of the test (AERA et al., 2014).

With this in mind and making use of the possibilities offered by the IRT, the severity hierarchies of the BAPQ subscales were compared across test versions in the Aloof dimension (the only one presenting unidimensionality across versions). As two items of this subscale showed invariant functioning, the lack of invariance hindered the comparison of the resulting severity hierarchies of both Aloof subscales between the English and Spanish versions. For explaining why Aloof severity hierarchies functioned differently between BAPQ versions, we referred to "genuine cultural specifics" (Hambleton, 1994) and Rogler's three-level framework (1993) exposed in Chapter One (cultural aspects may mediate the endorsement of behaviours, their self-perception and subsequent rating in scales). It seems that the BAP expressions are susceptible to being modulated by cultural aspects. Such is the influence of the culture in the formation of every psychological phenomenon from its very beginning that the effect of its influence resulted uncontrollable and unpredictable. This is even more relevant in light of the dimensional perspective of autism as BAP expressions could not only differ in severity, but also qualitatively.

Nevertheless, adding the differential functioning of the Aloof subscale to precedent unfavourable evidence regarding the internal structure of the BAPQ, we concluded that we cannot discern whether the BAP expression is modulated by an external condition or that the measurement tool needs a reconfiguration. This is the reason why it is so important to take these considerations as prerequisites when a test is being developed, trying to create adequate, reliable, and stable items across test applications and subsequent adaptations to different languages. The implications of this lack of invariant functioning are clear not only because the interpretations of the scores are challenged, but also since some authors are interested in studying the influence of cultural aspects in the development of autism and the expressions of autistic behaviours in different countries.

7.2 BAP tests Misfunctioning Led to Construction of the Broad Autism Phenotype—International Test (BAP-IT)

As exposed in the working plan, once adverse evidence about the BAPQ was disclosed, it was time for the development of a new international test for the assessment of the dimensional structure of the BAP across two domains which comprehend seven subdomains corresponding to the DSM-V proposal for ASD (APA, 2013). With this purpose in mind, in Study 3, we tried to reassign the items of the AQ, BAPQ, and SRS-II to one of those seven autistic subdomains divided into two main dimensions and we explored which contents of these tests would suit this two-dimensional operationalization of the BAP. This step was mandatory since, as reviewed in Chapter One, none of the tests replicated the two-dimensional structure of ASD/BAP (see Table 1, Chapter One, Block One). Empirical evidence from Study 3 arose from parallel analysis which suggested that the items of the three tests could be grasped into two sources of variance. These results supported the two-dimensional BAP operationalization.

Once empirical evidence was provided and complementary to it, in the second place, counting on the help of a group of experts in ASD/BAP, we evaluated whether the items of those tests were relevant for assessing the specific subdomain they had been assigned to, and if the seven BAP subdomains were correctly represented by the resulting selected items. After having selected the most relevant items inside each subdomain, experts concluded that the three tests failed in capturing important aspects of the BAP. Furthermore, two of the four subdomains of Patterns of restricted and repetitive behaviour were not presented in the final pool of items and, thus, the BAP was not adequately represented by them. Therefore, we suggested that new items should be created to fill the theoretical gaps in the BAP structure. Apart from these conclusions that have been thoroughly exposed in Chapter Five (Block Two), we consider that the main highlight of Study 3 was that the resulting pool of items did not get enough support for representing the new measurement tool of the two-dimensional operationalisation of the BAP.

On the other hand, we wanted to know experts' opinions about the proposed dimensional operationalization of the BAP. All considered, we believe the theoretical foundations for the updated BAP operationalization have been settled. Beyond the implications of these outcomes for this dissertation, the reconceptualization of the BAP, its insertion in the autism severity continuum, and its structure represent a milestone in the field of autism. The implications and derivations of this have been slightly commented on in Chapter One (Block One). Finally, the transcendence for further studies will be stated further down (see Chapter Eight).

In direct confluence with previous adverse evidence from Studies 1 and 2, conclusions from Study 3 encouraged us to take a step forward and construct a new tool for the assessment of the two-dimensional structure of the BAP. This, which had been preliminarily considered in the working plan as the worst possible scenario, has become the main objective of the present research and the BAP-IT its main outcome. The BAP-IT arrived to fill some gaps in the measurement of the lower levels of autism in Spanish and English-speaker countries. The test is formed by 20 items distributed in two main dimensions and it has tackled several utility issues observed in previous tests. One of its main strengths is that it is the only test targeting the two-dimensions structure of the BAP in line with the autism continuum hypothesis. The other main strength of the BAP-IT is that its items have been empirically chosen simultaneously in English and Spanish, emerging parallelly in two countries with different languages and cultures. These two advantages may be the base of the expansion of the international research on ASD and BAP since the BAP-IT allows reconnecting ASD/BAP concepts empirically and opens the door to study them cross-culturally. Furthermore, its Spanish and English versions will ease the conduction of subsequent adaptations of the test to all the countries where Spanish and English are the official languages and to other languages as well.

In addition to this, we can also highlight that the BAP-IT has amended previous concerns in the psychometric assessment of the BAP addressed in Chapter One (Block One). This amendment comes from a thorough process of content construction always trying to be faithful to the proposed autism structure and content in the DSM-V (APA, 2013) and following a theory-driven approach from the beginning to the end of the BAP-IT development. Experts' depuration of the pool of items also helped to amend previous concerns.

The timeline of the research should be resumed to underline the natural path of this last study. Primarily, as exposed, the BAP-IT aimed to represent the first international test of the two-dimensional structure of the BAP (aligned with the ASD operationalisation in the DSM-V, and considering the most recent research in the field) and, following a theory-driven approach, we attempted that the final test contemplated items of every BAP subdomain. Secondly, we provided validity evidence based on the internal structure of the BAP and about its invariance between the Spanish and the English versions. Thirdly, validity evidence based on subscales' relationships to other variables was supplied. Directly linked to this, BAP-IT subscales' connections to other variables also talked in favour of the two-dimensional structure of the BAP which was compared with the previous three-domain structure of the BAP targeted in the BAPQ.

The results of Study 4 complement the outcomes of Study 3 and provide further evidence of the BAP operationalization and dimensional structure. In Study 4, we also addressed problems regarding the lack of relevant content that have been documented in previous tests (e.g., AQ, BAPQ). We opted for a theory-driven approach due to its multiple advantages, which have been exposed in Chapter One. Under this premise, we did not stop the research when achieved adequate indicators, but instead, we tried to recompose the sample of items adding alternative ones from Study 3 (see Chapter 5 Block Two) which filled the content gaps in the BAP structure. As those items had been denoted as relevant for a specific autistic subdomain, we thought that they would enrich the BAP-IT test content. After all, we think that the meticulous procedure we have followed in fourth study has contributed to the fact that all the BAP subdomains were represented in the final test with at least two items each.

Trying to maintain a stable test structure, only unidimensional items were simultaneously selected in the two countries. Furthermore, and what is more relevant, this solution resulted in fully metric-invariance between the Spanish groups and partially metric-invariance with respect to the English sample. This suggests the steadiness of the BAP-IT structure will be maintained even cross-culturally. Directly connected with that, both BAP-IT subscales have shown good internal consistency which is also relevant considering that previous literature has reported diverse substructures inside each autistic domain (Georgiades et al., 2007; Lecavalier, 2005; Miranda-Linne & Melin, 2002; Tadevosyan-Leyfer et al., 2003; Van Lang et al., 2006). Finally, the fact that the three subdomains of the SOCIAL-BAP and the four ones of the RIRE-BAP variate under the same dimension have supported the internal structure of the BAP-IT following the operationalization of the BAP and, by extension, the ASD one in the DSM-V (APA, 2013).

Apart from that, outcomes from the internal structure of the BAP-IT have also uncovered further important findings, namely the considerable reduction in the number of items, an issue that deserves deeper discussion. One main implication of the last study has been the observed relationships between specific subdomains from different domains. As commented in Study 4 (see Chapter Six, Block Two), some items, especially those of SB, were sharing variance with RIRE-BAP and some items of RC were with SOCIAL-BAP (they showed cross-loadings). As a particular example, items referring to nonverbal communicative behaviours (e.g., poorly integrated communication, abnormalities in eye contact and body language, or problems to understand and use gestures) have shared variance with items from RA which refers to echolalia and idiosyncratic speech (repetitive and restricted speech). After deleting items presenting cross-loadings, non-verbal communicative behaviours were represented in SB items and stereotyped speech ones were still in RA. This represents support for previous literature in the topic (see, Golysheva, 2019; Kanner, 1943; Rutter, 1978; Sterponi & Shankey, 2012) and supports the ASD structure proposed in the DSM-V (APA, 2013) which we consider a valuable contribution of this research, particularly taking into account that this finding was expected after the outcomes of Studies 1 and 2.

Concluding with Study4, our last objective was to gather validity evidence based on dimensions' relationships to other variables (AERA et al., 2014). BAP-IT subscales correlations with BAPQ ones were as expected. Subsequent zero-order correlations with the subscales of the SRS-II and RCS also provided evidence about the good representation of the central aspects of all the autistic subdomains inside either SOCIAL-BAP or RIRE-BAP.

On the other hand, as a part of this last objective, we tested whether the proposed two-dimensional structure of the BAP-IT, besides being more parsimonious, was able to predict better the variance of the subscale scores of the related variables than the three-factor model of the BAPQ (more similar to autism disorder in the DSM-IV [APA, 1994]). After exploring the partial correlations between the nodes of the networks in conjunction with zero-order correlations among the subscales' scores, we drew some important highlights. First of all, the patterns of relationships commented in Study 4 gave us reasons for concluding that SOCIAL-BAP may be better capturing the variance of core behaviours related to the first dimension of the ASD (Impaired social communication and social relation; Criterion A; APA, 2013) than Aloof which could more focused on motivational engagement, disinterest in or lack of enjoyment of social interactions (Hurley et al., 2007). The core social deficits that were not connected to Aloof were related to the Pragmatic Language which may be, by contrast, reflecting more social aspects than it was originally intended (Hurley et al., 2007). Apart from that, we think that Pragmatic Language connections to RIRE-BAP and Rigid also suggested that it may be not only considering further social impairments but, also, some stereotyped patterns of speech.

Regarding the connections of RIRE-BAP mainly with RRB (SRS-II), we could conclude that RIRE-BAP better explained the variance of all the central aspects of the second dimension of autism (Restricted and repetitive patterns of behaviours; Criterion B; APA, 2013) including aspects of RB and RC subdomains (restricted behaviours) as zero-order correlations have suggested. Yet, relevantly, it seems to also comprehend items that refer to RA and RD subdomains that were underrepresented in the previous test such as the BAPQ as experts have highlighted in Study 3 (see Chapter 3, Block 2). Taken alone, Rigid from the BAPQ seems to be more focused on assessing the core rigid aspects of autism (routine seeking and extreme discomfort about changes; Hurley et al., 2007). Similarly, Aloof has been the second subscale with the highest partial correlation with Routine Seeking (RCS). From these results, we inferred that the three subscales of the BAPQ were more linked to sameness while the BAP-IT ones capture better the variance attributable to most of the core aspects of autism. Highlights of Study 3 complement and support this conclusion. Directly connected with this, it seems that Study 4 has addressed the main limitations of Study 3 as the parsimonious two-dimensional structure of the BAP-IT has been able to comprehend key autistic behaviours that experts have reported missing in the three-domain structure of the BAPQ.

After discussing these findings and how they tipped the balance in favour of SOCIAL-BAP and RIRE-BAP, we consider it relevant to delve into the specific case of pragmatic language deficits. As the reader may note, one of the most significant changes of ASD in the DSM-V (APA, 2013) was to collapse the three autistic domains into two and relocate the behaviours referred to in Deficits in the pragmatic language domain into the remaining two ones. Unlike the BAPQ, with the BAP-IT we have committed to this dimensional conceptualization of autism. Zero-order correlations of Pragmatic Language subscale (BAPQ) with both SOCIAL-BAP and RIRE-BAP and with Social Communication and RRB subscales of the SRS-II support the idea that the content of pragmatic language domain has not been eliminated from the ASD structure, and has entirely collapsed with the social impairments domain as some authors have suggested. What we could conclude is that social and non-verbal deficits of communication form now part of both dimensions of the ASD structure where the social aspect of non-verbal communication (e.g., poorly coordinated gestures or facial expressions with speech) would be included in the first dimension (Impaired social communication and social relation; Criterion A; APA, 2013) and stereotyped and idiosyncratic speech in the second one (Restricted and repetitive patterns of behaviours; Criterion B; APA, 2013).

Finally, the main conclusion of this last study was that the BAP-IT better captures the variance attributable to most of the central behaviours of autism as it was intended. It seems that the two-dimensional structure of the BAP proposed in the BAP-IT is not only more parsimonious but also completely faithful to the current and worldwide accepted ASD operationalization (APA, 2013).

Altogether, the four empirical studies of this dissertation have yielded the Spanish version of a commonly used test for assessing the BAP, the BAPQ, and a new measurement tool, the BAP-IT, in two languages, English and Spanish. Furthermore, varied validity evidence has been collected: both theoretically (literature review, expert judgment) and empirically (i.e., common factor methods, structural equation models, and empirical networks) for exploring the two-dimensions of the BAP operationalization, the BAP-IT internal structure, and its relations with other variables. Comparisons between Spanish and English versions of the test subscales or internal structures have also been disclosed. Unlike precedent BAP tests, we have followed a theory-driven approach for developing the BAP-IT and, after all our efforts, finally, we have presented a test with a stable and invariant structure also representative of all the core autistic behaviours at BAP levels of severity. With it, we have overcome a utility issue in the field and we have opened the door to multiple new possibilities that will be briefly commented on in the next chapter.

CHAPTER EIGHT. FUTURE DIRECTIONS AND WRITER'S FINAL APPRECIATIONS

In this last chapter, chapter eight, the future work that this dissertation has opened for the coming years will be exposed summarily. Furthermore, I will provide my personal appreciation, based on the formation I have received during the past five years, about the most important aspects of my work, its implications and how I have personally experienced the process.

8.1 Future Directions

Once we have reunited and discussed the main findings of the four studies separately and as a sum, we would like to provide some further directions before putting an end to this last chapter of my dissertation. Furthermore, we think this chapter becomes essential since it symbolizes the prospective path of my research career, the direct research lines we have opened to continue investigating about the BAP, the ones we have been working on but, due to time limitation, it was not possible to include in this dissertation, and the new promising opportunities my research has prepared for me. Each of those is going to be further commented and an analytical approach is going to be proposed.

8.1.1 Exploratory attempt to outline the nomological net of the BAP

Following a sequential order, the first path we should discuss is the unfinished research we have been working on for the last five months. Once evidence based on the internal structure of the BAP-IT was provided, the psychometric properties of our test have been studied and validity evidence based on subscales scores' relations with other variables has been recollected from two different approaches (i.e., correlations between subscales and empirical network analysis). The last approach applied in Study 4, empirical network analysis, enabled us to compare the predictive capability of the BAP-IT subscale scores and against the BAPQ ones.

This last approach could bring many advantages to the BAP-IT and, hence, we consider that further validity evidence should be recollected. Concerning this, plenty of former literature has reported the relationships of the core autistic behaviours at BAP levels with different cognitive processes, behaviours, or personality expressions. For instance, Baron-Cohen (2002) and Wheelwright et al. (2006) studies have proposed that behaviours at BAP levels may involve a trend for activities of systematizing (i.e., to analyse, understand, predict, control, and construct rule-based systems) to the detriment of those requiring empathizing (i.e., to identify other person's emotions and thoughts and respond to them accordingly). These studies have enriched the BAP operationalization as they have located the phenotype in a wider nomological net (Cronbach & Meehl, 1955). From them, it can be supposed that the SOCIAL-BAP will be linked to problems in social aspects of the communication (pragmatic use of the language), social cognition procedures such as empathy or emotional facial recognition. In the same way, as SOCIAL-BAP also involves distress and dislike for social events and a preference for solitude, these behaviours are supposed to be equally connected to this dimension. While slightly less studied, restricted and repetitive behaviours (RIRE-BAP) have been equally observed in conjunction with conditions of extreme worry about changes or distress about unexpected alterations in daily routines and resistance to trying new things. Likewise, several psychological conditions, particularly, social phobia and obsessive-compulsive disorder (OCD) have been reported to be comorbid to autism (e.g., Bolton et al., 1999; Micali et al., 2004; Wainer et al., 2011)

In an exploratory *Study 5*, further validity evidence based on SOCIAL-BAP and RIRE-BAP scores' relations to other variables was gathered. A sample of 268 university students completed the BAP-IT and four additional questionnaires. In particular, the subscales of the BAP-IT (SOCIAL-BAP and RIRE-BAP) were studied in relation to the subscales of the Spanish versions of the BAPQ (Aloof, Pragmatic Language, Rigid; Hurley et al., 2007; see Chapters Three and Four, Block Two), the Resistance to Change Scale ([RCS], Routine Seeking, Short-term focus, Emotional Reaction, and Cognitive Rigidity; Oreg, 2003; Arciniega & González, 2009), the Interpersonal Reactivity Index ([IRI], Fantasy, Perspective Taking, Personal Distress, and Empathic Concern; Davis, 1980; Pérez-Albéniz et al., 2003), the Obsessive-Compulsive Inventory-Revised ([OCI-R], Hoarding, Checking, Ordering, Mental Neutralizing, Washing, and Obsessing; Foa et al., 2002; Fullana et al, 2005) and the Social Phobia Inventory ([SPIN], Social Phobia overall score; Connor et al., 2000; García-López et al., 2010). Some of those tests had been included in Study 4 (i.e., BAPQ, RCS) and given that their relationships to either SOCIAL-BAP or RIRE-BAP met expectations and that the contents their subscales refer to are closely connected to autistic central behaviours as discussed in Study 4, we decided to include them in this new network. We hypothesised that SOCIAL-BAP would be positive and moderate-to-high related to Aloof (BAPQ) and SPIN overall (SPIN), moderate and positive correlated with Pragmatic Language (BAPQ), Rigid (BAPQ), and the subscales of the IRI (negative in the case of Empathic Concern and Perspective Taking). On its behalf, RIRE-BAP would be highly positively high correlated with Rigid (BAPQ) and with the subscales of the OCI-R. Finally, taking into consideration Study 4 highlights, we expected to find moderate and positive connections with RCS subscales.

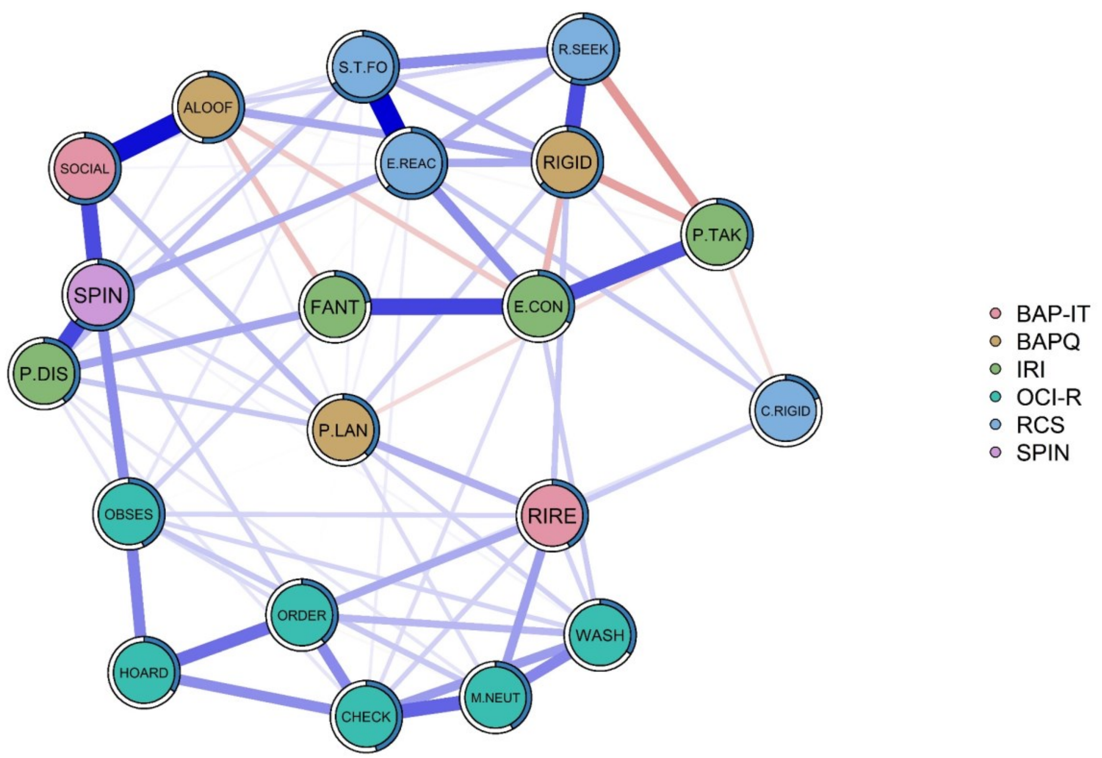
The results of this exploratory study are going to be briefly exposed. However, it may be noted that these results are still preliminary as the sample is being collected at the moment. Preliminary zero-order correlations (not presented here) suggest that SOCIAL-BAP highest correlations are with Aloof (BAPQ) and SPIN overall score. This subscale is also moderately related to Routine Seeking (RCS), Short-term focus (RCS), Emotional Reaction (RCS), and Pragmatic Language (BAPQ) and Rigid (BAPQ). RIRE-BAP highest correlations are with Mental Neutralizing and Ordering (OCI-R) followed by moderate correlations with the rest of the OCI-R subscales (Hoarding, Checking, Neutralizing, Washing, and Obsessing) and Pragmatic Language (BAPQ) and Rigid (BAPQ).

Following the approach in Study 4, empirical network models have been applied to explore the partial correlations between variables and to compare the predictive capability of SOCIAL-BAP and RIRE-BAP raw scores (BAP-IT subscales) with BAPQ raw scores (Aloof, Pragmatic Language, and Rigid subscales). As in Study 4, we are also trying to explore the predictive capability of the BAP-IT subscales in comparison to the BAPQ ones. The empirical partial connections present in the network (see Figure 1) suggest that SOCIAL-BAP is more connected to Aloof (BAPQ), SPIN overall (SPIN), and Pragmatic Language (BAPQ). RIRE-BAP is more related to Mental Neutralizing (OCI-R), Ordering (OCI-R), Pragmatic Language (BAPQ), Rigid (BAP), Cognitive Rigidity (RCS), Checking (OCI-R), Obsessions (OCI-R), and Washing (OCI-R). Regarding the BAPQ, Aloof (BAPQ) presents linkings to SOCIAL-BAP, Rigid

(BAPQ), Fantasy (negative partial correlation; IRI), and Empathic Concern (negative partial correlation; IRI). Pragmatic Language (BAPQ) to RIRE-BAP, SOCIAL-BAP, Personal Distress (IRI), Rigid (BAPQ), Washing (OCI-R), Perspective Taking (negative partial correlation; IRI), and SPIN overall score; and, Rigid (BAPQ) to Routine Seeking (RCS), Perspective Taking (negative partial correlation; IRI), Aloof (BAPQ), Emotional Reaction (IRI), Short-term Focus (RCS), Empathic Concern (negative partial correlation; IRI), RIRE-BAP, Pragmatic Language (BAPQ), and Cognitive Rigidity (RCS). SOCIAL-BAP is connected indirectly to Personal Distress (IRI) and Obsessions (OCI-R) via SPIN, and to Rigid (BAPQ) and Fantasy (IRI) through Aloof (BAPQ). RIRE-BAP has also shown indirect connections to SOCIAL-BAP (as mentioned) and Personal Distress (IRI) via Pragmatic Language (BAPQ); and to Routine Seeking (RCS) and Emotional Reaction (RCS) through Rigid (BAPQ).

Figure 1

Empirical Network Model of the BAP-IT, BAPQ, IRI, OCI-R, RCS, and SPIN Subscales



Note. Nodes are coloured regarding their test as displayed in the legend.
BAP-IT subscales: SOCIAL (SOCIAL-BAP), RIRE (RIRE-BAP); BAPQ subscales: ALOOF, P.LAN (Pragmatic Language), RIGID; IRI subscales: E.CON (Empathic Concern), FANT (Fantasy), P.TAK (Perspective Taking), P.DIS (Personal Distress); SPIN: SPIN (SPIN overall score); RCS subscales: R.SEEK (Routine Seeking), E.REAC (Emotional Reaction), S.T.FO (Short-term Focus), C.RIGID

(Cognitive Rigidity); OCI-R subscales: HOARD (Hoarding), CHECK (Checking), ORDER (Ordering), M.NEUT (Mental Neutralizing), WASH (Washing), OBSES (Obsessions).

As the reader may note, the exploratory nature of this analysis and the lower number of participants only from Spain make the exposed results only informative and that they should be taken as a possible promising guide for further studies with the BAP-IT. Having said that, some preliminary results seem so promising that they are going to be highlighted. In the first place, the relationships between SOCIAL-BAP with Aloof (BAPQ), Pragmatic Language (BAPQ), and SPIN should be mentioned. The relationships with the first two subscales have been as expected and matched with results from Study 4, supporting the conclusion that SOCIAL-BAP is assessing a wide range of social central behaviours of autism. SOCIAL-BAP and SPIN correlation has been also congruent with previous research (e.g., Bellini, 2004; De Bruni et al., 2007; Farrugia & Hudson, 2006). As the relationship is positive, it implied that higher Social Phobia behaviours (e.g., persistent, intense fear or anxiety about a specific social situation; APA, 2013) could be similar to SC behaviours (Deficits in developing, maintaining, and understanding relations) such as management of anxiety and negative thoughts in social contexts and preference for loneliness. By contrast, it is worth mentioning that Aloof (BAPQ) connection with SPIN has been very small. It might be possible that those anxiety expressions in social contexts are not included in the Aloof subscale and that this subscale comprehends other social behaviours such as shared-social enjoyment, social reciprocity behaviours, and prosocial behaviours (Hurley et al., 2007) as was suggested in Study 3 and Study 4. Finally, expected relations to IRI subscales have hardly been found. SOCIAL-BAP is indirectly related to Fantasy while Aloof is negatively connected to Fantasy and Empathic Concern (IRI). Personal Distress (IRI) is more connected to Pragmatic Language (BAPQ) and Perspective Taking (IRI) to Rigid (BAPQ). As argued above, once a larger Spanish community sample will have been collected and complemented with an English one, those results could vary. Were this not to happen, an alternative test or explanation should be explored.

Regarding RIRE-BAP, its connection with Rigid (BAPQ) has been congruent with Study 4 and so it has been its higher connection with Cognitive Rigidity (RCS) and with the rest of the subscales of the RCS. Its connections to OCI-R subscales have also been congruent with precedent literature which has suggested that OCD behaviours could be similar to restricted and repetitive behaviours ones that are highly representative and core of autism (e.g., Bolton et al., 1998; Micali et al., 2004; Wainer et al., 2011). More particularly, Compulsion subscales (Washing, Ordering, and Checking; APA, 2013) could be linked to almost all RA behaviours (e.g., stereotyped motor movements, ordering, or echolalia) and some RD behaviours such as excessive smelling and palpation. Furthermore, Obsessions and Mental Neutralizing (OCD Obsession dimension; APA, 2013) could be associated with RC behaviours (e.g., excessive, restricted, and perseverant interests, strong attachment or preoccupation with unusual objects).

Thus, as mentioned above, more participants are needed to obtain more stable results and draw clearer conclusions. Yet, we cannot disregard the plentiful advantages that empirical networks models offer to the study of the comorbid conditions of autism. Findings derived from this approach would represent an extremely valuable contribution to research concerning the whole autism phenomenon and not only for more intense levels of expression.

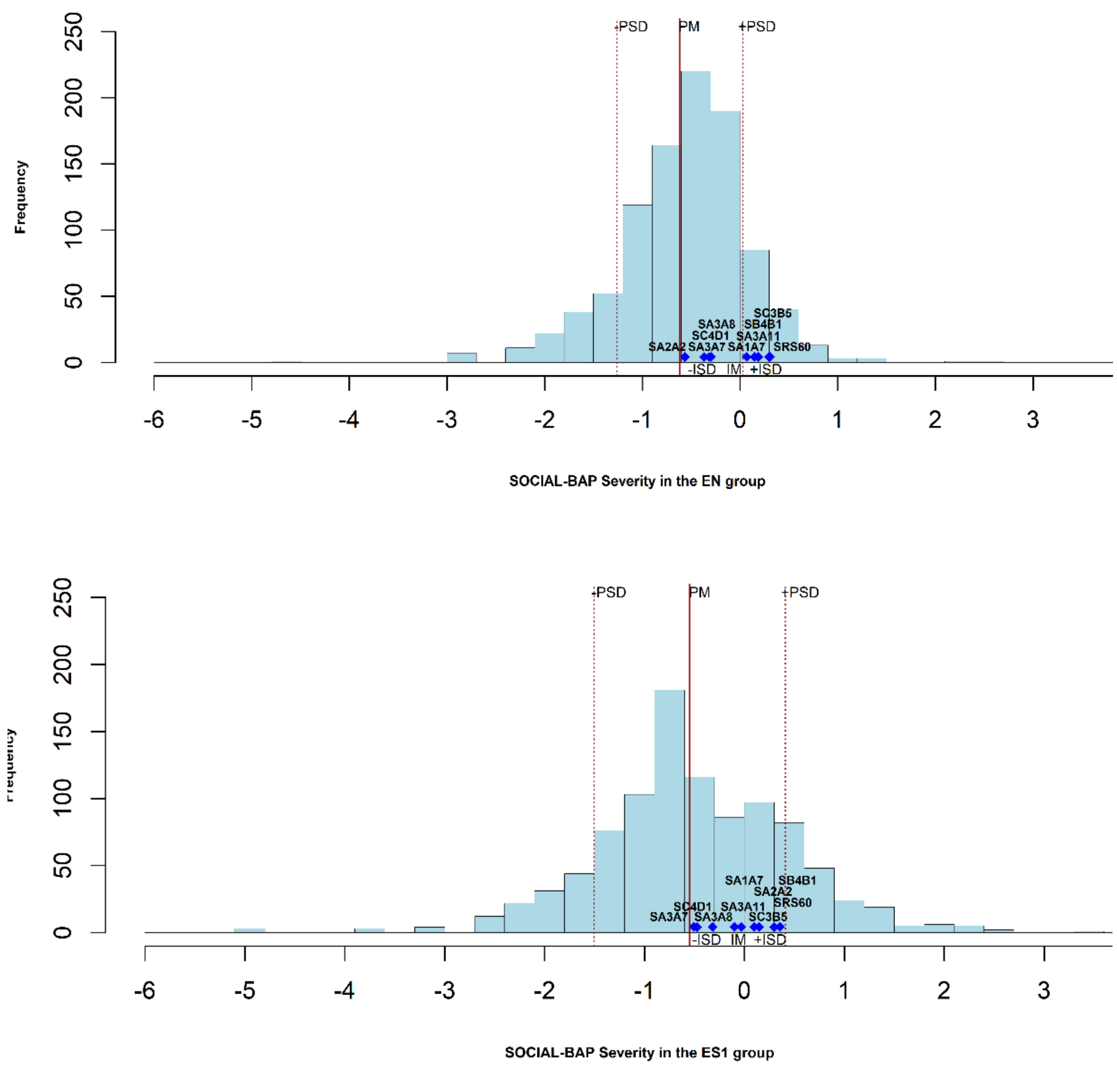
8.1.2 Exploratory attempt to map the severity of the BAP-IT contents

As it happened with Studies 1 and 2, once validity evidence around the BAP-IT has been provided, we need to review the BAP operationalization proposed in the BAP-IT with a dimensional severity approach regarding both items and persons. We refer to Rasch analysis. Indeed, this analysis should have been included in the present dissertation not as a complementary approach but as an essential one. If we have a last look at the theoretical framework of this research and the theories and approaches, we founded the research in, we can observe that: the severity dimensional approach of ASD has been extensively commented on, the simplified autism severity continuums have been drawn, the BAP has been re-operationalized according to a two-dimensional structure and the BAP-IT has been developed to study it. However, the dimensional structure of the BAP-IT is not fully understood with the evidence provided in Study 4 since the severity hierarchies of SOCIAL-BAP and RIRE-BAP have not been analyzed either has the differential functioning of the subscales or specific items between test versions. The fact that those results have not been introduced does not mean the study was not conducted and analysis was not performed.

In this regard, a preliminary Study 6 exposes the results of this analysis. Even if it is still unfinished work, the results from this study report that the items from both SOCIAL-BAP and RIRE-BAP are susceptible to being scaled according to their BAP/TEA severity in SOCIAL-BAP and RIRE-BAP. Additionally, this time items are tapping levels of severity more similar to those observed in the general population (compared to the BAPQ-SP in Study 2). An example of each SOCIAL-BAP severity hierarchy either in the Spanish or English version is displayed in Figure 2. Further analysis also informs about the unidimensionality of SOCIAL-BAP and RIRE-BAP, provides supportive evidence to Study 4, and also talks about the adequateness of the items' estimators (infit and outfit) and dimensions' reliability.

Figure 2

An Example of the Scalability of the English and Spanish versions of the SOCIAL-BAP Dimension



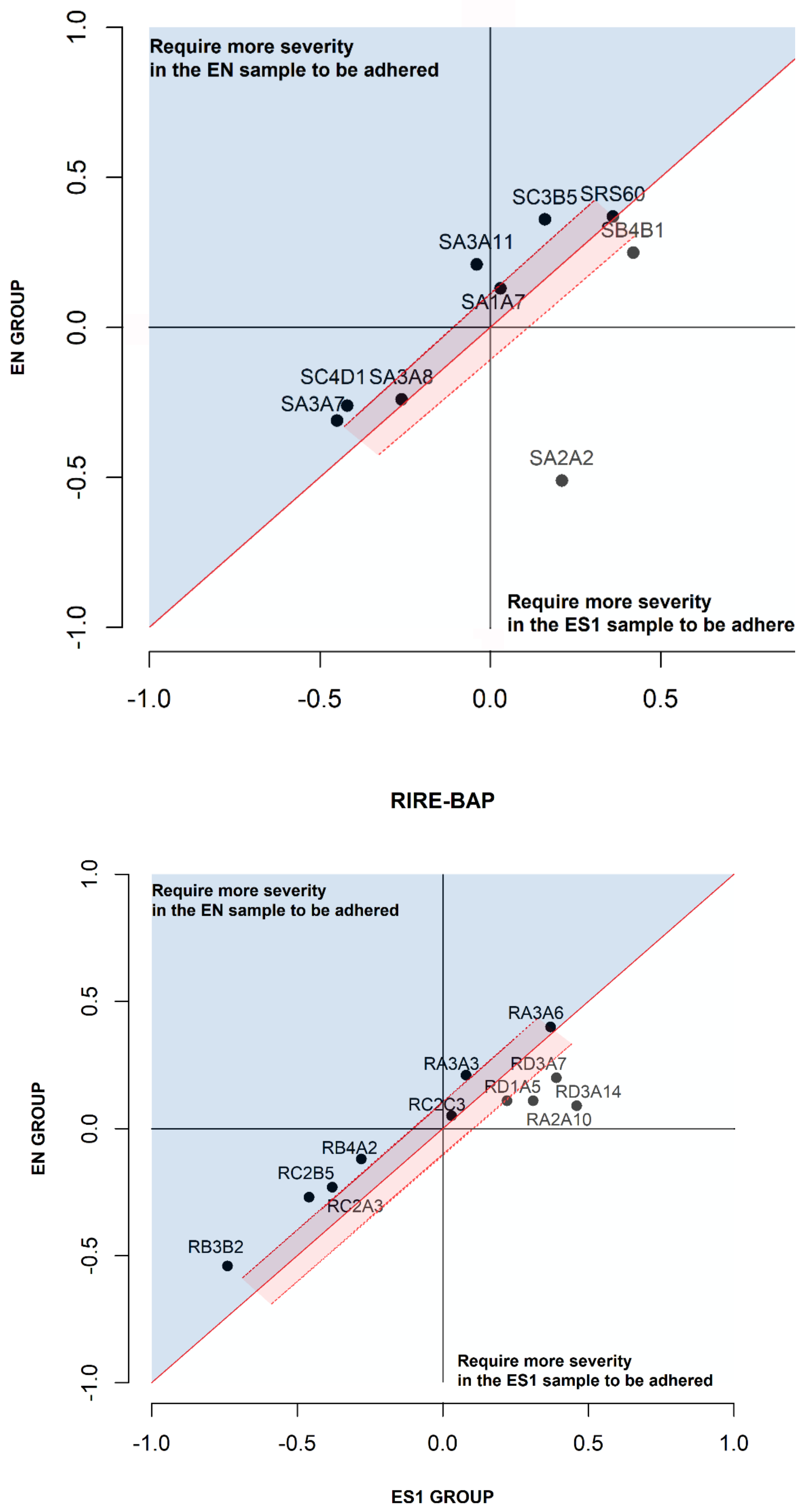
Note. Histograms represent the distribution of the person scores regarding either EN or ES1 groups. Blue diamonds placed on the x-axis represent item measures along the same continuum as person scores. Item labels were distributed near their respective diamond to avoid text overlapping[18].
PM = person mean score; +PSD = person standard deviation above the person mean; -PSD = person standard deviation below the person mean; IM = item mean; +ISD = item standard deviation above the item mean; -ISD = item standard deviation below the item mean.

Finally, one of the most important aspects under analysis was differential test functioning (DFT). The performance between both groups in SOCIAL-BAP and RIRE-BAP as depicted by their items (DFT) was compared with identity plots (Bond & Fox, 2015). An example of identity plots for SOCIAL-BAP and RIRE-BAP can be seen in Figure 3. The order of the severity hierarchies of SOCIAL-BAP and RIRE-BAP (DFT) differ between the Spanish and English versions being the SOCIAL-BAP the most susceptible dimension to cultural idiosyncrasies (the highest differential severity of the items is observed inside SOCIAL-BAP dimension).

[18] English SOCIAL-BAP severity hierarchy was (from less severity to more severity): SA2A2, SA3A7, SC4D1, SA3A8, SA1A7, SA3A11, SB4B1, SC3B5, SRS60. Spanish SOCIAL-BAP severity hierarchy was (from less severity to more severity): SA3A7, SC4D1, SA3A8, SA3A11, SA1A7, SC3B5, SA2A2, SRS60, SB4B1.

Figure 3

Comparison of SOCIAL-BAP and RIRE-BAP Severity Hierarchies Between the English and Spanish Groups

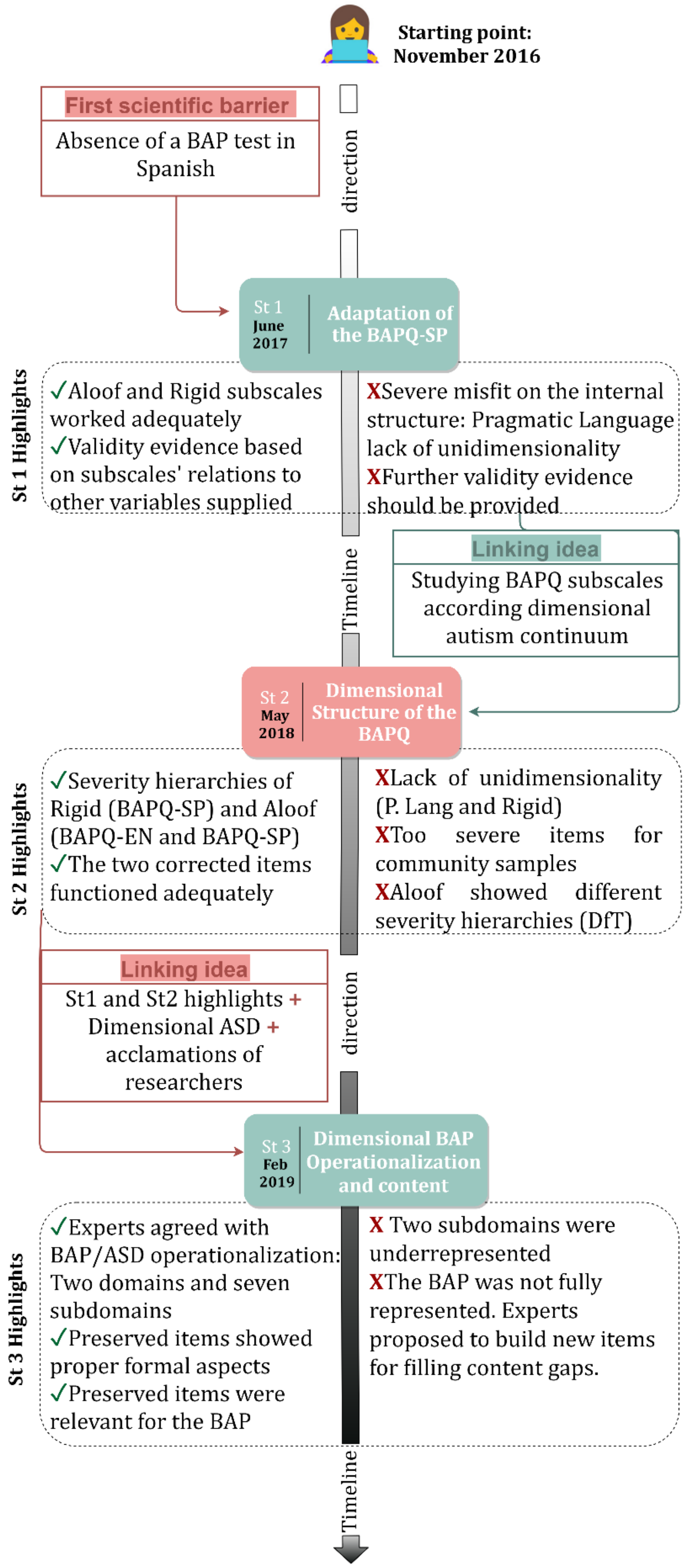


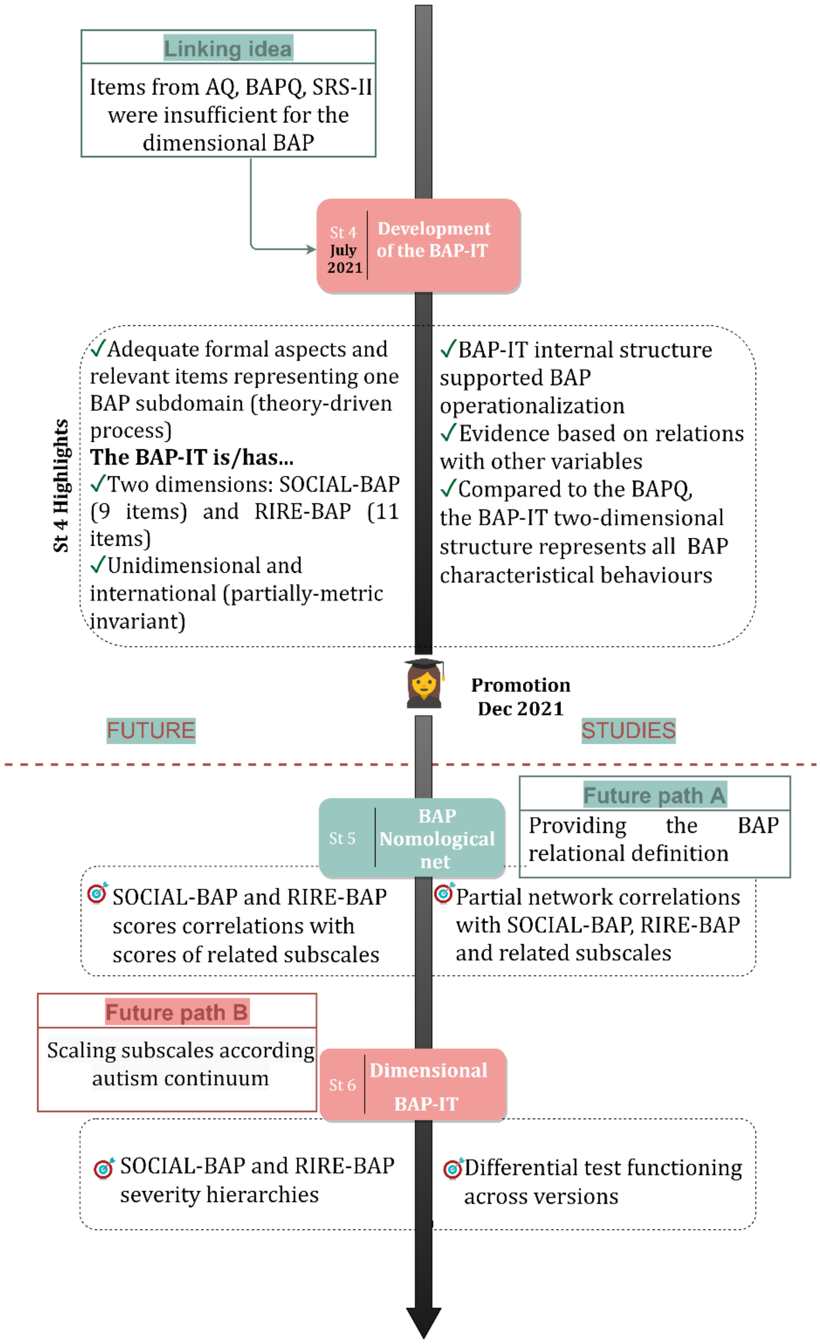
Note. These scatterplots represent item severity of each subscale version in each axis; the x-axis represents always the item values of the Spanish group and the y-axis of the English one. We added an identity line (in red) symbolizing the expected values for severity invariance and a pair of 95% quality control lines (red dotted lines).

Although a throughout inspection of the results is needed and they should be interpreted in the context of the newest research in BAP, Study 6 will close a stage in this research, "the dimensional conceptualization and international assessment of the BAP". Finally, it may be useful to conclude this section by depicting the main outcomes of the present dissertation covering from Study 1 to these upcoming Studies 5 and 6 as well as the links between one study and the next one (see Figure 4).

Figure 4

A Representation of the Timeline of my Research and a Summary of the Main Conclusions





Note. The diagram shows the timeline of my research sorted by studies. Each study is represented by a coloured label (either blue or pink) inside which the reader can find the numeral order of the study (i.e., St 1, St 2, St 3, St 4, St 5 or St 6), the date and a summary of the study purpose. The main conclusions of each study have been attached to it (framed in dotted lines). Studies have been chained including the reasons which link one study to the following. Two exploratory studies of my research have been proposed (St 5 and 6).

8.1.3 *Further directions*

Once this stage of the present dissertation is closed, a complementary investigation could be prepared. One research path should include further evidence based on the BAP-IT relations with other variables and comorbid disorders such as attention deficit hyperactivity disorder, OCD or social communication disorder. This path will connect the BAP and ASD with other related and updated disorders in the DSM-V and will insert it inside a bigger net. The empirical network analysis is the perfect candidate to conduct this kind of analysis which will provide a new and more comprehensive vision of psychological disorders.

Considering the dimensional severity approach, the natural path of this research will be to apply the BAP-IT to other samples with higher levels of severity such as first-degree relatives of people diagnosed with ASD. In this kind of study, it will be explored, for the first time, whether the two-dimensional structure of the test is replicated and, then, whether the items will be adequate to assess the levels of severity presented in those other samples or more severe items should be included. The severity hierarchies of both BAP-IT dimensions will be provided, being the IRT the best approach in this kind of study. This research path will help deepen into the differential expressions of the same psychological processes related to autism at BAP levels of severity and will provide extremely valuable information to the severity continuums of autism.

Relatedly, we cannot omit former studies which laid the foundations of the BAP as a psychological phenomenon (see Folstein & Rutter, 1977). Those studies suggested a vulnerability for ASD that could be inherited in "moderate-in-severity" expressions of the prototypical ASD behaviours which were named as BAP. Following studies reasoning, whether the BAP represents a genetic liability for autism observed in relatives of people diagnosed with ASD, BAP behaviours would not be observed in relatives of people diagnosed with other psychological conditions such as Down Syndrome. Accordingly, the BAPQ was designed and applied in two samples; the first sample comprehended non-autistic parents of children diagnosed with ASD and the second one parents of children diagnosed with Down Syndrome. The performance of both samples was compared and subscales cut-offs were established. Concerning this, some authors could be interested in continuing exploring the genetic predisposition of ASD and, to that end, it could be interesting to apply the BAP-IT and compare the performance of parents of children diagnosed with ASD to parents of children diagnosed with another childhood psychological condition.

Complementary to this, and a mandatory process in any further application and adaptation of the BAP-IT to other languages or versions (e.g., extensive version, informant-report version, paper-based version), will be to replicate the two-dimensional structure of the test, providing evidence-based on its internal structure and external net and further psychometric properties. Equally important, cross-cultural invariance should be studied any time the test will be adapted to another language. All the evidence derived from these studies, which will not be necessarily conducted by the authors of the test, will contribute to enriching the study of autism severity continuums and reveal how different aspects of the context or the culture would affect the assessment process.

Finally but not least, once Studies 5 and 6 will be completed, it will also enable the study of the differential condition of these psychological phenomena at different levels of the autism continuums such as BAP levels or ASD ones. Thus, we or other researchers interested in the topic will be capable of conducting, at last, complex experiments that are difficult to perform in people with ASD levels of severity (as exposed in Chapter One) in more accessible and easy samples such as those expressing BAP levels of severity. As the BAP and the ASD are parts of the same continuum of severity and share expressions, inferences made on BAP levels could be extended to ASD ones (e. g., De Groot & Van Strien, 2017; Rubenstein & Chawla, 2018). As the reader can perceive, the present research has offered plenty of derivations and applications. Further implications of having accessibility to a test that assesses the BAP have been extensively exposed in Chapter One, Block One and could be equally applied here, as the BAP-IT has resulted to be an adequate test with countless potential uses.

8.2. Writer's final appreciations

On the 21st of October of 2021, I finished writing the main conclusions of my research. On the 22nd of October of 2021, I browsed the scientific data webs for the last time. In this last search, I entered the same keywords I have been using during the past five years: "autism", "autistic traits", "broad autism phenotype", "autism and the general population", "milder autistic traits", and so on. Two examples of the results of this last search are displayed in Table 1.

Table 1

Findings of Each Keyword Sorted by Data Source

Data source	Number of results	Number of results
	"Broad autism phenotype"	"Autism" AND "General population"
Google Scholar	2,850	188,000
Last five years	1,210	20,400
Last year	235	7,590
ProQuest	1,039	1,368
Last five years	490	572
Last year	116	106
PubMed	552	15,764
Last five years	252	9,180
Last year	49	2,257
Scopus	414	12,320
Last five years	176	6,731
Last year	38	1,658
Web of Science	185	1,740
Last five years	90	870
Last year	16	129
APA PsycInfo	5	1,368
Last five years	-	540
Last year	-	170
PSICODOC	3	52
Last five years	-	8
Last year	-	3

My first perception of these results was that the interest in the BAP or autistic expressions at general population levels has recently increased. If we pay attention to the number of publications in the last five years, we can observe that they have been growing, mostly internationally and in English. A quick inspection of the titles and abstracts of the research papers (once duplicate papers would have been deleted) suggests that the majority of these results represents applied research; that is, authors applied a well-known test for further research purposes such as studying the social cognition deficits observed in adults with BAP levels of severity compared to adults with lower levels of severity, without testing the psychometric properties or the internal structure of the test they have applied. Normally, there should not be any problem in this regard. However, just one hour of searching about the best BAP tests is enough to realise that the situation is, at least, confusing and that authors should act carefully when deepening in this topic.

The high number of results also makes me wonder why no author has tried to adapt any BAP test to Spanish earlier. Mostly bearing in mind the huge number of test adaptations to other languages (see Chapter One, Block One). Even if the scientific field was not interested in adapting a BAP test into Spanish, if papers about the BAP are increasing, it implies that researchers are interested in the topic. In this regard, when we read an article after the arrival of the DSM-V in 2013, normally, researchers try to define the BAP as the phenotypical expression of ASD and even try to compare and equate the expressions of both conditions. In parallel, the gap in the conceptualization between both is omitted most of the time. The rest of the time, authors mislaid the operationalization of the BAP and defined it similarly to ASD but without providing empirical evidence in this regard or applying tests with a different internal structure than the one suggested theoretically.

Directly branched with this issue is the scope of my research. <u>Study 1</u>, the first published study, has already 659 accesses, eight citations and the BAPQ-SP has been applied in four studies for further research purposes. <u>Study 3,</u> the second published study, has already 214 access. Yet, it is not as important the interest in my research as it is the implications of each study. <u>Study 1</u> together with <u>Study 2</u> served to provide evidence to the validity of the unquestionable-until-then internal structure of the BAPQ. Adverse evidence based on it has put the BAPQ in the spotlight and has helped to further investigate the internal structure of the test, the outdated operationalization of the BAP and the existent wedge between what researchers claim theoretically and what empirical evidence points to. <u>Study 3</u> has represented the first study in reviewing the content comprehended in the most applied BAP tests, the AQ, the BAPQ, and the SRS-II, and whether it represents all the relevant autistic central behaviours and not others. We performed this study in the present dissertation since we considered it an essential and preliminary step to the use of any measurement tool. It has helped with the theory-driven approach we followed in <u>Study 4</u> as it helped to locate the main aspect of autism that should be included in either SOCIAL-BAP or RIRE-BAP.

Afterwards, it comes the turn to present what I considered the masterpiece of the present dissertation, <u>Study 4</u>. It represents all the empirical support to the two-dimensional operationalization of the BAP we have fiercelessly defended throughout the present dissertation mainly with the help of well-known experts in the field of autism and with previous studies. It has also materialized it in an international test, the BAP-IT. With it, as it has been previously exposed, we have overcome a utility issue in the field of autism and we have opened the door to countless possibilities, some of them already mentioned above. Finally, it also may be noted that even if the SRS-II counts on a validated Spanish version, this is the first time that an autism test has been simultaneously developed in two different countries. This, which may be partially shadowed in the conclusions of Study 4, leads to promising future studies in the international assessment of the BAP/ASD, considering the mostly plurilingual and globalized world where we live at the present.

By the time I was finishing my research, I gave a thorough thought to my theoretical approach as I think this is something every researcher needs to do before submitting him/her dissertation. Evidence from the literature and empirical studies was in line with our theoretical approach and has also supported the inferences made on the scores of the two domains of the BAP-IT. This has helped me deal with my doubts about some parts of my research or about the theoretical framework we have defended. However, this compulsory feeling of searching for contradictory clues has not stopped until that last search on the Internet I have presented at the beginning of the section. Eventually, I can firmly conclude that I could not change my perspective to date. Even if it is true that this is an arena still under construction, I think the contribution of the present research to the field of autism is going to be pretty discussed and at the same time highly valuable. Likewise, I also believe that the future paths of it will be multiple and it will even help other authors continue with their research.

On the other hand, I cannot disregard the promising opportunities my dissertation has brought. We cannot forget that the research career is a matter of connections and opportunities. In this sense, the PhD Degree not only settles the foundations of an adequate, faithful, and professional researcher but offers plenty of opportunities for collaborations representing a chain where all the links are interconnected. I cannot disrespect either that the original purpose of my formation was to adapt a test to the Spanish to conduct further experimental studies about the

BAP and related psychological processes like facial emotional processing and the BAP-IT has opened the door to all these kinds of studies. After suggesting future routes to continue with my research, it seems to me that my research career has been already programmed for another five years.

To conclude, I would love to devote this last paragraph of my dissertation to the people who got directly involved with it. I could not conclude my work without making a tiny special tribute to my supervisors and lab partners because, without their enormous help, this dissertation would have resulted in an almost impossible endeavour. The quality of my research should not be evaluated in isolation since what I have presented here is just a mere reflection of the great formation, I have received during the past five years. I have started this manuscript explaining the three areas of psychology that were going to be involved in my PhD formation and I would like to finish highlighting the luck I have had and the importance of interdisciplinary work. Likewise, it was not only the knowledge they transmitted to me but mainly, the direct work and the time they have spent to give birth to this dissertation. There would not be any better way than closing this section by expressing my utmost gratitude to my supervisors for giving me the opportunity of taking part in an interdisciplinary group of great professionals.

REFERENCES

List of References

A

Abubakar, A., Ssewanyana, D., de Vries, P. J., & Newton, C. R. (2016). Autism spectrum disorders in sub-Saharan Africa. *The Lancet Psychiatry, 3*(9), 800-802. https://doi.org/10.1016/S2215-0366(16)30138-9

Adolphs, R., Spezio, M. L., Parlier, M., & Piven, J. (2008). Distinct face-processing strategies in parents of autistic children. *Current Biology*, *18*(14), 1090-1093. https://doi.org/10.1016/j.cub.2008.06.073

Alarcón RD. (2009). Culture, cultural factors and psychiatric diagnosis: review and projections*. World Psychiatry 8,* 131–9. http://dx.doi.org/10.1002/j.2051-5545.2009.tb00233.x.

Amaral, D., Geschwind, D., & Dawson, G. (Eds.). (2011). *Autism spectrum disorders*. Oxford University Press.

American Educational Research Association. (2014). *Standards for educational and psychological testing*. American Educational Research Association [AERA]. American Psychological Association [APA]. National Council on Measurement in Education [NCME].

American Psychiatric Association. (1952). *Diagnostic and statistical manual of mental disorders (DSM-I®)*. American Psychiatric Pub.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders (DSM-III®)*. American Psychiatric Pub.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (DSM-IV®)*. American Psychiatric Pub.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV-TR®)*. American Psychiatric Pub.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*(4), 561-573. https://doi.org/10.1007/BF02293814

Arciniega, L. M., & González, L. (2009). Validation of the Spanish-language version of the resistance to change scale. *Personality and Individual Differences*, *46*(2), 178-182. https://doi.org/10.1016/j.paid.2008.09.024

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences, 38*, 451–460. https://doi.org/10.1016/j.paid.2004.04.022

B

Bailey, A., Palferman, S., Heavey, L., & Le Couteur, A. (1998). Autism: The phenotype in relatives. *Journal of autism and developmental disorders*, *28*(5), 369-392.https://doi.org/10.1023/A:1026048320785

Bakare, M. O., Tunde-Ayinmode, M. F., Adewuya, A. O., Bello-Mojeed, M. A., Sale, S., James, B. O., ... & Orovwigho, A. O. (2015). Recognition of Autism Spectrum Disorder (ASD) symptoms and knowledge about some other aspects of ASD among final year medical students in Nigeria, Sub-Saharan Africa. *BMC Research Notes, 8*, 1-8. https://doi.org/10.1186/s13104-015-1433-0

Baker, J. P. (2013). Autism at 70--redrawing the boundaries. *The New England journal of medicine*, *369*(12), 1089. https://doi.org/10.1056/NEJMp1306380

Bandim, J. M., Ventura, L. O., Miller, M. T., Almeida, H. C., & Costa, A. E. S. (2003). Autism and Möbius sequence: an exploratory study of children in northeastern Brazil. *Arquivos de Neuro-psiquiatria*, *61*(2A), 181-185.

Barker, E. T., Hartley, S. L., Seltzer, M. M., Floyd, F. J., Greenberg, J. S., & Orsmond, G. I. (2011). Trajectories of emotional well-being in mothers of adolescents and adults with autism. *Developmental Psychology, 47*(2), 551–561. https://doi.org/10.1037/a0021268

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in cognitive sciences*, *6*(6), 248-254. https://doi.org/10.1016/S1364-6613(02)01904-6

Baron-Cohen, S., & Hammer, J. (1997). Is autism an extreme form of the" male brain"? *Advances in Infancy research*, *11*, 193-218.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of autism and developmental disorders*, *31*, 5-17. https://doi.org/10.1023/A:1005653411471

Basha, S. (2014). First national public opinion survey: Palestinian's knowledge and understanding of autism, 2014. *Italian Journal of Special Education for Inclusion, 2*, 87-96.

Basu, S. N., Kollu, R., & Banerjee-Basu, S. (2009). AutDB: a gene reference resource for autism research. *Nucleic acids research*, *37*, D832-D836. https://doi.org/10.1093/nar/gkn835

Bernier, R., Gerdts, J., Munson, J., Dawson, G., & Estes, A. (2012). Evidence for broader autism phenotype characteristics in parents from multiple-incidence autism families. *Autism Research*, *5*, 13-20.https://doi.org/10.1002/aur.226

Berry, J. W. (1992). Acculturation and adaptation in a new society. *International migration*, *30*, 69-69.

Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain research*, *1380*, 42-77.

Bishop, D. V. (2003). *The children's communication checklist: CCC-2*. Harcourt Assessment.

Bishop, D. V., Maybery, M., Wong, D., Maley, A., Hill, W., & Hallmayer, J. (2004). Are phonological processing deficits part of the broad autism phenotype? *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *128*, 54-60. https://doi.org/10.1002/ajmg.b.30039

Bölte, S., Poustka, F., & Constantino, J. N. (2008). Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Research*, *1*(6), 354-363. https://doi.org/10.1002/aur.49

Bölte, S., Westerwald, E., Holtmann, M., Freitag, C., & Poustka, F. (2011). Autistic traits and autism spectrum disorders: The clinical validity of two measures presuming a continuum of social communication skills. *Journal of autism and developmental disorders, 41*, 66-72. https://doi.org/10.1007/s10803-010-1024-9

Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., ... & Rutter, M. (1994). A case-control family history study of autism. *Journal of child Psychology and Psychiatry, 35*(5), 877-900. https://doi.org/10.1111/j.1469-7610.1994.tb02300.x

Bolton, P. F., Pickles, A., Murphy, M., & Rutter, M. (1998). Autism, affective and other psychiatric disorders: patterns of familial aggregation. *Psychological medicine*, *28*(2), 385-395. https://doi.org/10.1017/S0033291797006004

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). L. Erlbaum.

Bordini, D., Lowenthal, R., Gadelha, A., Araujo, G. M. D., Mari, J. D. J., & Paula, C. S. (2015). Impact of training in autism for primary care providers: a pilot study. *Brazilian Journal of Psychiatry, 37*, 63-66. https://doi.org/10.1590/1516-4446-2014-1367

Broderick, N., Wade, J. L., Meyer, J. P., Hull, M., & Reeve, R. E. (2015). Model invariance across genders of the broad autism phenotype questionnaire. *Journal of autism and developmental disorders*, *45*(10), 3133-3147. https://doi.org/10.1007/s10803-015-2472-z

Brown, T. A. (2015). Confirmatory factor analysis for applied research. Guilford publications.

Bruni, T. P. (2014). Test review: Social responsiveness scale–Second edition (SRS-2). *Journal of Psychoeducational Assessment*, *32*(4), 365-369. https://doi.org/10.1177/0734282913517525

Brunsdon, V. E., Colvert, E., Ames, C., Garnett, T., Gillan, N., Hallett, V., ... & Happé, F. (2015). Exploring the cognitive features in children with autism spectrum disorder, their co-twins, and typically developing children within a population-based sample. *Journal of Child Psychology and Psychiatry*, *56*(8), 893-902. https://doi.org/10.1111/jcpp.12362

Brunsdon, V. E., & Happé, F. (2014). Exploring the 'fractionation' of autism at the cognitive level. *Autism*, *18*(1), 17-30. https://doi.org/10.1177/1362361313499456

C

Camodeca, A., & Voelker, S. (2016). Automatic and controlled processing and the Broad Autism Phenotype. *Psychiatry Research*, *235*, 169–176. https://doi.org/10.1016/j.psychres.2015.11.012

Campbell, D. G., Reichle, N. C., & Van Bourgondien, M. E. (1996). The autism survey: An evaluation of reliability and validity. *Journal of autism and developmental disorders*, *26*(6), 621-634. https://doi.org/10.1007/BF02172351

Caron, K. G., Schaaf, R. C., Benevides, T. W., Gal, E. (2012). Cross-cultural comparison of sensory behaviours in children with autism. *Am J Occup Ther, 66,* e77– 80. http://dx.doi.org/10.5014/ajot.2012.004226

Carretero-Dios, Hugo, & Pérez, Cristino (2007). Standards for the development and review of instrumental studies: Considerations about test selection in psychological research. *International Journal of Clinical and Health Psychology, 7*(3), 863-882. [fecha de Consulta 2 de abril de 2021]. ISSN: 1697-2600. Disponible en: https://www.redalyc.org/articulo.oa?id=33770319

Chakraborty, R. (2017). Configural, metric and scalar invariance measurement of academic delay of gratification scale. *Int. J. Human. Soc. Stud, 3*(3), 1-9. http://dx.doi.org/10.20908/ijarsijhss.v3i3.7903

Chien, Y. L., Wu, Y. Y., Chiu, Y. N., Liu, S. K., Tsai, W. C., Lin, P. I., ... & Chien, W. H. (2011). Association study of the CNS patterning genes and autism in Han Chinese in Taiwan. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(6), 1512-1517. https://doi.org/10.1016/j.pnpbp.2011.04.010

Chouinard, P. A., Unwin, K. L., Landry, O., & Sperandio, I. (2016). Susceptibility to optical illusions varies as a function of the autism-spectrum quotient but not in ways predicted by local–global biases. *Journal of Autism and Developmental Disorders, 46*, 2224–2239. https://doi.org/10.1007/s10803-016-2753-1

Chung, K-M., Jung, W., Yang, J., Ben-Itzchak, E., Zachor, D. A., Furniss, F., Heyes, K., Matson, J. L., Matson, J., Kozlowski, A., Barker, A. (2012). Cross cultural differences in challenging behaviours of children with autism spectrum disorders: An international examination between Israel, South Korea, the United Kingdom, and the United States of America. *Res Autism Spectr Disord, 6*, 881–9. http://dx.doi.org/10.1016/j.rasd.2011.03.016.

Cohen, D. J., & Volkmar, F. R. (1997). *Handbook of autism and pervasive developmental disorders*. John Wiley & Sons Inc.

Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S.R., Woodhouse, E., …, Bolton, P. (2015). Heritability of Autism spectrum disorder in a UK population-based twin sample. *JAMA Psychiatry*, pp. 1-9. http://dx.doi.org/10.1016/10.1001/jamapsychiatry.2014.3028

Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Weisler, R. H., & Foa, E. (2000). Psychometric properties of the Social Phobia Inventory (SPIN): New self-rating scale. *The British Journal of Psychiatry*, *176*(4), 379-386. https://doi.org/10.1192/bjp.176.4.379

Constantino, J. N. (2011). The quantitative nature of autistic social impairment. *Pediatric research*, *69*(8), 55-62. https://doi.org/10.1203/PDR.0b013e318212ec6e

Constantino, J. N. (2018). Deconstructing autism: from unitary syndrome to contributory developmental endophenotypes. *International Review of Psychiatry*, *30*(1), 18-24. https://doi.org/10.1080/09540261.2018.1433133

Constantino, J. N., & Charman, T. (2016). Diagnosis of autism spectrum disorder: reconciling the syndrome, its diverse origins, and variation in expression. *The Lancet Neurology*, *15*(3), 279-291. https://doi.org/10.1016/S1474-4422(15)00151-9

Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L., ... & Reich, W. (2003). Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of autism and developmental disorders*, *33*(4), 427-433. https://doi.org/10.1023/A:1025014929212

Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale: SRS-2*. Torrance, CA: Western Psychological Services.

Constantino, J. N., & Gruber, C. P. (2017). *Spanish version of the social responsiveness scale: SRS-2*. Torrance, CA: Western Psychological Services.

Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry, 45*(4), 719-726. https://doi.org/10.1111/j.1469-7610.2004.00266.x

Constantino, J. N., Lajonchere, C., Lutz, M., Gray, T., Abbacchi, A., McKenna, K., ... & Todd, R. D. (2006). Autistic social impairment in the siblings of children with pervasive developmental disorders. *American Journal of Psychiatry*, *163*(2), 294-296. https://doi.org/10.1176/appi.ajp.163.2.294

Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders.

Journal of Developmental and Behavioral Pediatrics, 21, 2–11. https://doi.org/10.1097/00004703-200002000-00001

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Archives of general psychiatry, 60*, 524-530. http://dx.doi.org/10.1001/archpsyc.60.5.524

Constantino, J. N., Zhang, Y. I., Frazier, T., Abbacchi, A. M., & Law, P. (2010). Sibling recurrence and the genetic epidemiology of autism. *American Journal of Psychiatry*, *167*(11), 1349-1356. https://doi.org/10.1176/appi.ajp.2010.09101470

Coonrod EE, Stone WL. (2004). Early concerns of parents of children with autistic and nonautistic disorders. *Infants Young Child, 17*, 258–68.

Cribb, S. J., Olaithe, M., Di Lorenzo, R., Dunlop, P. D., & Maybery, M. T. (2016). Embedded figures test performance in the broader autism phenotype: a meta-analysis. *Journal of Autism and Developmental Disorders*, *46*(9), 2924-2939. https://doi.org/10.1007/s10803-016-2832-3

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL 32887.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, *52*(4), 281. https://doi.org/10.1037/h0040957

Cruz, L. P., Camargos-Júnior, W., & Rocha, F. L. (2013). The broad autism phenotype in parents of individuals with autism: A systematic review of the literature. *Trends in Psychiatry and Psychotherapy*, *35*(4), 252-263. https://doi.org/10.1590/2237-6089-2013-0019

Cuccaro, M. L., Wright, H. H., Rownd, C. V., Abramson, R. K., Waller, J., & Fender, D. (1996). Brief report: Professional perceptions of children with developmental difficulties: The influence of race and socioeconomic status. *Journal of autism and developmental disorders*, *26*(4), 461-469. https://doi.org/10.1007/BF02172830

D

Daley TC. (2002). The need for cross-cultural research on the pervasive developmental disorders. *Transcult Psychiatry*, *39*, 531–50. https://doi.org/10.1177/136346150203900409

Davids, R. C., Groen, Y., Berg, I. J., Tucha, O. M., & van Balkom, I. D. (2016). Executive functions in older adults with autism spectrum disorder: Objective performance and subjective complaints. *Journal of Autism and Developmental Disorders*, *46*(9), 2859-2873. https://doi.org/10.1007/s10803-016-2831-4

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.

Dawson, G., Estes, A., Munson, J., Schellenberg, G., Bernier, R., & Abbott, R. (2007). Quantitative assessment of autism symptom-related traits in probands and parents: Broader phenotype autism symptom scale. *Journal of Autism and Developmental Disorders, 37*(3), 523–536. https://doi.org/10.1007/s10803-006-0182-2.

Delgado-Rico, E., Carretero-Dios, H., & Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology Espana*, *12*(3), 449-460. https://doi.org/10.5167/uzh-64551

De Groot, K., & Van Strien, J. W. (2017). Evidence for a broad autism phenotype. *Advances in Neurodevelopmental Disorders*, *1*(3), 129-140. https://doi.org/10.1007/s41252-017-0021-9

De Leeuw, A., Happé, F., & Hoekstra, R. A. (2020). A conceptual framework for understanding the cultural and contextual factors on autism across the globe. *Autism Research, 13*(7), 1029-1050. https://doi.org/10.1002/aur.2276

De Vries, P. J. (2016). Thinking globally to meet local needs: autism spectrum disorders in Africa and other low-resource environments. *Current Opinion in Neurology, 29*(2), 130-136. https://doi.org/10.1097/WCO.0000000000000297

Devlin, B. & Scherer, S. W. (2012). Genetic architecture in autism spectrum disorder. *Current Opinion in Genetics & Development, 22*(3). pp. 229-237. https://doi.org/10.1016/j.gde.2012.03.002

DiCriscio, A. S., & Troiani, V. (2018). The broader autism phenotype and visual perception in children. *Journal of autism and developmental disorders*, *48*(8), 2809-2820. https://doi.org/10.1007/s10803-018-3534-9

D'Cruz, A. M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., & Sweeney, J. A. (2013). Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology*, *27*(2), 152. https://doi.org/10.1037/a0031721

Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., ... & Shih, A. (2015). Autism screening and diagnosis in low resource settings: Challenges and opportunities to enhance research and services worldwide. *Autism Research, 8*(5), 473-476. https://doi.org/10.1002/aur.1575

E

Elsabbagh, M., Divan, G., Koh, Y. J., Kim, Y. S., Kauchali, S., Marcín, C., ... & Fombonne, E. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism research*, *5*(3), 160-179. https://doi.org/10.1002/aur.239

Embretson, S. E. & McCollam, K. M. S. (2000). Psychometric approaches to understanding and measuring intelligence. In R.J. Sternberg (Ed.). *Handbook of intelligence* (pp. 423-444). Cambridge University Press.

English, M. C., Gignac, G. E., Visser, T. A., Whitehouse, A. J., & Maybery, M. T. (2019). A comprehensive psychometric analysis of autism-spectrum quotient factor models using two large samples: Model recommendations and the influence of divergent traits on total-scale scores. *Autism Research.* https://doi.org/10.1002/aur.2198

Eriksson, K. (2013). Autism-spectrum traits predict humor styles in the general population. *Humor, 26*(3), 461-475. https://doi.org/10.1515/humor-2013-0030

F

Faso, D. J., Corretti, C. A., Ackerman, R. A., & Sasson, N. J. (2016). The broad autism phenotype predicts relationship outcomes in newly formed college roommates. *Autism*, *20*(4), 412-424. https://doi.org/10.1177/1362361315585733

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science, 8(*4), 370-378. https://doi.org/10.1177/1948550617693063

Floris, D. L., Lai, M. C., Auer, T., Lombardo, M. V., Ecker, C., Chakrabarti, B., ... & Suckling, J. (2016). Atypically rightward cerebral asymmetry in male adults with autism stratifies individuals with and without language delay. *Human Brain Mapping*, *37*, 230-253. https://doi.org/10.1002/hbm.23023

Fisher, J. E., Guha, A., Heller, W., & Miller, G. A. (2020). Extreme-groups designs in studies of dimensional phenomena: Advantages, caveats, and recommendations. *Journal of Abnormal Psychology, 129*(1), 14–20. https://doi.org/10.1037/abn0000480

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: development and validation of a short version. *Psychological assessment*, *14*(4), 485. https://doi.org/10.1037/1040-3590.14.4.485

Folstein, S., & Rutter, M. (1977). Infantile autism: A genetic study of 21 twin pairs. *Journal of Child Psychology and Psychiatry, 18*, 297–321. https://doi.org/10.1111/j.1469-7610.1977.tb00443.x

Frazier, T. W., Embacher, R., Tilot, A. K., Koenig, K., Mester, J., & Eng, C. (2015). Molecular and phenotypic abnormalities in individuals with germline heterozygous PTEN mutations and autism. *Molecular psychiatry*, *20*(9), 1132-1138.https://doi.org/10.1038/mp.2014.125

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2014). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism*, *18*(1), 31-44. https://doi.org/10.1177/1362361313500382

Frazier, T. W., Youngstrom, E. A., Hardan, A. Y., Georgiades, S., Constantino, J. N., & Eng, C. (2015). Quantitative autism symptom patterns recapitulate differential mechanisms of genetic transmission in single and multiple incidence families. *Molecular autism*, *6*(1), 1-12. https://doi.org/10.1186/s13229-015-0050-z

Frazier, T. W., Youngstrom, E. A., Sinclair, L., Kubu, C. S., Law, P., Rezai, A., ... & Eng, C. (2010). Autism spectrum disorders as a qualitatively distinct category from typical behavior in a large, clinically ascertained sample. *Assessment*, *17*(3), 308-320. https://doi.org/10.1177/1073191109356534

Freeth, M., Sheppard, E., Ramachandran, R., Milne, E. (2013). A cross-cultural comparison of autistic traits in the UK, India and Malaysia. *J Autism Dev Disord*, *43,* 2569–83. http://dx.doi.org/10.1007/s10803-013-1808-9

Freitag, C. M., Kleser, C., Schneider, M., & von Gontard, A. (2007). Quantitative assessment of neuromotor function in adolescents with high functioning autism and Asperger syndrome. *Journal of autism and developmental disorders, 37(*5), 948-959. http://dx.doi.org/10.1007/s10803-006-0235-6

Fullana, M. A., Tortella-Feliu, M., Caseras, X., Andión, Ó., Torrubia, R., & Mataix-Cols, D. (2005). Psychometric properties of the Spanish version of the Obsessive–Compulsive Inventory—Revised in a non-clinical sample. *Journal of anxiety disorders*, *19*(8), 893-903. https://doi.org/10.1016/j.janxdis.2004.10.004

G

Gable, R.K., & Wolf, J.W. (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings.* Boston: Kluwer Academic.

Garcia-Lopez, L. J., Bermejo, R. M., & Hidalgo, M. D. (2010). The Social Phobia Inventory: Screening and cross-cultural validation in Spanish adolescents. *The Spanish Journal of Psychology*, *13*(2), 970-980. https://doi.org/10.1017/S1138741600002614

Gau, S. S. F., Lee, C. M., Lai, M. C., Chiu, Y. N., Huang, Y. F., Kao, J. D., & Wu, Y. Y. (2011). Psychometric properties of the Chinese version of the Social Communication Questionnaire. *Research in Autism Spectrum Disorders, 5*(2), 809-818. https://doi.org/10.1016/j.rasd.2010.09.010

Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., Lee, A. B., et al. (2014). Most genetic risk for autism resides with common variation. *Nature Genetics*, *46*(8), 881–885. https://doi.org/10.1038/ng.3039.

Gravetter, F. J., & Forzano, L. A. B. (2009). *Research methods for psychology*. Cengage Learning Australia.

Geisinger, K. F. (2016). Test development and validation. In J. C. Norcross, G. R. VandenBos, D. K. Freedheim, & B. O. Olatunji (Eds.), *APA handbook of clinical psychology: Theory and research* (p. 393–405). American Psychological Association. https://doi.org/10.1037/14773-016

Georgiades, S., Szatmari, P., Boyle, M., Hanna, S., Duku, E., Zwaigenbaum, L., ... & Pathways in ASD Study Team. (2013). Investigating phenotypic heterogeneity in children with autism spectrum disorder: a factor mixture modelling approach. *Journal of Child Psychology and Psychiatry*, *54*(2), 206-215. https://doi.org/10.1111/j.1469-7610.2012.02588.x

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., Roberts, W., ... & Mahoney, W. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(2), 188-196. https://doi.org/10.1097/01.chi.0000242236.90763.7f

Gerdts, J. (2013). *An examination of the broader autism phenotype in simplex and multiplex families* (Doctoral dissertation). http://hdl.handle.net/1773/21786

Gerdts, J., & Bernier, R. (2011). The broader autism phenotype and its implications on the aetiology and treatment of autism spectrum disorders. *Autism Research and Treatment*, *2011*, 545901 https://doi.org/10.1155/2011/545901.

Gerdts, J., Bernier, R., Dawson, G., & Estes, A. (2013). The broader autism phenotype in simplex and multiplex families. *Journal of Autism and Developmental Disorders, 43*(7), 1597–1605. https:// doi.org/10.1007/s10803-012-1706-6

Geschwind, D. H. (2009). Advances in autism. *Annual review of medicine*, *60*, 367-380. https://doi.org/10.1146/annurev.med.60.053107.121225

Glumbić, N., & Brojčin, B. (2012). Factor structure of the Serbian version of the Children's Communication Checklist-2. *Research in Developmental Disabilities, 33*(5), 1352-1359. https://doi.org/10.1016/j.ridd.2012.03.010

Golysheva, M. D. (2019). A Review on Echolalia in Childhood Autism. *Advances in Social Science, Education and Humanities Research, 333*, 200-204. https://doi.org/10.1007/s40489-015-0067-4

Gotham, K., Pickles, A., & Lord, C. (2012). Trajectories of autism severity in children using standardized ADOS scores. *Pediatrics*, *130*(5), e1278-e1284. https://doi.org/10.1542/peds.2011-3668

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *American journal of psychiatry*, *160*(4), 636-645. https://doi.org/10.1176/appi.ajp.160.4.636

Grinter, E. J., Van Beek, P. L., Maybery, M. T., & Badcock, D. R. (2009). Brief report: Visuospatial analysis and self-rated autistic-like traits. *Journal of Autism and*

Developmental Disorders, *39*(4), 670-677. https://doi.org/10.1007/s10803-008-0658-3

H

Hahler, E.-M., & Elsabbagh, M. (2014). Autism: A global perspective. *Current Developmental Disorders Reports, 2*, 58–64. https://doi.org/10.1007/s40474-014-0033-3

Haladyna, T. M. (2004). *Developing and validating multiple-choice test items*. Routledge.

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., ... & Risch, N. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of general psychiatry*, *68*(11), 1095-1102. https://doi.org/10.1001/archgenpsychiatry.2011.76

Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment, 10*(3), 229-244.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (Eds.). (2004). *Adapting educational and psychological tests for cross-cultural assessment*. Psychology Press.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Sage Publications.

Happé, F., Briskman, J., & Frith, U. (2001). Exploring the cognitive phenotype of autism: weak "central coherence" in parents and siblings of children with autism: I. Experimental tests. *Journal of child psychology and psychiatry*, *42*(3), 299-307. https://doi.org/10.1111/1469-7610.00723

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature neuroscience*, *9*(10), 1218-1220. https://doi.org/10.1038/nn1770

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. Psychological Assessment, 7(3), 238–247. https://doi.org/10.1037/1040-3590.7.3.238

Helland, W. A., Biringer, E. V. A., Helland, T., & Heimann, M. (2009). The usability of a Norwegian adaptation of the Children's Communication Checklist Second

Edition (CCC-2) in differentiating between language impaired and non-language impaired 6-to 12-year-olds. *Scandinavian Journal of Psychology, 50*(3), 287-292. https://doi.org/10.1111/j.1467-9450.2009.00718.x

Hernandez, A., Hidalgo, M. D., Hambleton, R. K., & Gomez-Benito, J. (2020). International test commission guidelines for test adaptation: A criterion checklist. *Psicothema*, *32*(3), 390-398. https://doi.org/10.7334/psicothema2019.306

Hilton, C. L., Zhang, Y., Whilte, M. R., Klohr, C. L., & Constantino, J. (2012). Motor impairment in sibling pairs concordant and discordant for autism spectrum disorders. *Autism*, *16*(4), 430-441. https://doi.org/10.1177/1362361311423018

Hirokawa, K., Kimura, T., Ikehara, S., Honjo, K., Ueda, K., Sato, T., & Iso, H. (2020). Associations Between Broader Autism Phenotype and Dietary Intake: A Cross-Sectional Study (Japan Environment & Children's Study). *Journal of autism and developmental disorders*, *50*(8), 2698-2709.https://doi.org/10.1007/s10803-020-04380-z

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. *Journal of autism and developmental disorders*, *38*(8), 1555-1566. https://doi.org/10.1007/s10803-008-0538-x

Hoekstra, R. A., Bartels, M., Verweij, C. J., & Boomsma, D. I. (2007). Heritability of autistic traits in the general population. *Archives of pediatrics & adolescent medicine, 161*(4), 372-377. http://dx.doi.org/10.1007/s10803-008-0538-x

Hoekstra, R. A., Vinkhuyzen, A. A., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., ... & Van Der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). *Journal of autism and developmental disorders*, *41*(5), 589-596. http://dx.doi.org/10.1007/s10803-010-1073-0

Hong, S. J., Valk, S. L., Di Martino, A., Milham, M. P., & Bernhardt, B. C. (2018). Multidimensional neuroanatomical subtyping of autism spectrum disorder. *Cerebral Cortex*, *28*(10), 3578-3588. https://doi.org/10.1093/cercor/bhx229

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, 45*(2), 212–229. https://doi.org/10.1111/j.1469-7610.2004.00215.x

Howlin, P., & Moss, P. (2012). Adults with autism spectrum disorders. *The Canadian Journal of Psychiatry, 57*(5), 275–283. https://doi.org/10.1177/070674371205700502

Howlin, P., Moss, P., Savage, S., Bolton, P., & Rutter, M. (2015). Outcomes in adult life among siblings of individuals with autism. *Journal of Autism and Developmental Disorders, 45*(3), pp. 707-718. https://doi.org/10.1007/s10803-014-2224-5

Hughes, C., Leboyer, M., & Bouvard, M. (1997). Executive function in parents of children with autism. *Psychological medicine*, *27*, 209-220. https://doi.org/10.1017/S0033291796004308

Hurley, R. S., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of autism and developmental disorders*, *37*(9), 1679-1690. http://dx.doi.org/10.1007/s10803-006-0299-3

Hurst, R. M., Mitchell, J. T., Kimbrel, N. A., Kwapil, T. K., & Nelson-Gray, R. O. (2007). Examination of the reliability and factor structure of the Autism Spectrum Quotient (AQ) in a non-clinical sample. *Personality and Individual Differences, 43*(7), 1938-1949. http://dx.doi.org/10.1016/j.paid.2007.06.012

I

Ingersoll, B. (2010). Broader autism phenotype and nonverbal sensitivity: Evidence for an association in the general population. *Journal of Autism and Developmental Disorders*, *40*(5), 590-598. http://dx.doi.org/10.1007/s10803-009-0907-0

Ingersoll, B., & Hambrick, D. Z. (2011). The relationship between the broader autism phenotype, child severity, and stress and depression in parents of children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*, 337-344. https://doi.org/10.1016/j.rasd.2010.04.017

Ingersoll, B., Hopwood, C. J., Wainer, A., & Donnellan, M. B. (2011). A comparison of three self-report measures of the broader autism phenotype in a non-clinical sample. *Journal of Autism and Developmental Disorders, 41*, 1646–1657. https://doi.org/10.1007/s10803-011-1192-2

Ingersoll, B., & Wainer, A. (2014). The broader autism phenotype. In F. R. Volkmar, R. Paul, S. J. Rogers, & K. A. Pelphrey (Eds). *Handbook of Autism and Pervasive Developmental Disorders* (4th Edition). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118911389.hautc02

International Test Commission. (2017). The ITC Guidelines for Translating and Adapting Tests (Second edition). [www.InTestCom.org]

Irwing, P., Booth, T., & Hughes, D. J. (2018). *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale, and Test Development.* John Wiley & Sons. https://doi.org/10.1002/9781118489772

J

James, R. J., Dubey, I., Smith, D., Ropar, D., & Tunney, R. J. (2016). The Latent Structure of Autistic Traits: A Taxometric, Latent Class and Latent Profile Analysis of the Adult Autism Spectrum Quotient. *J Autism Dev Disord 46*, 3712–3728. https://doi.org/10.1007/s10803-016-2897-z

Jamil, R., Gragg, M. N., & DePape, A. M. (2017). The broad autism phenotype: Implications for empathy and friendships in emerging adults. *Personality and Individual Differences*, *111*, 199-204. https://doi.org/10.1016/j.paid.2017.02.020

Jobe, L. E., & White, S. W. (2007). Loneliness, social relationships, and a broader autism phenotype in college students. *Personality and individual differences*, *42*(8), 1479-1489. https://doi.org/10.1016/j.paid.2006.10.021

K

Khan, N., & Hombarume, J. (1996). Levels of autistic behaviour among the mentally handicapped children in Zimbabwe. *The Central African journal of medicine*, *42*(2), 36-39.

Kasmini, K., & Zasmani, S. (1995). Asperger's syndrome: a report of two cases from Malaysia. *Singapore medical journal*, *36*(6), 641-643.

Kanne, S. M., Wang, J., & Christ, S. E. (2012). The Subthreshold Autism Trait Questionnaire (SATQ): Development of a brief self-report measure of subthreshold autism traits. *Journal of Autism and Developmental Disorders, 42*(5), 769-780. https://doi.org/10.1007/s10803-011-1308-8

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250.

Kim, H. U. (2012). Autism across cultures: Rethinking autism. *Disability & Society*, *27*(4), 535-545. https://doi.org/10.1080/09687599.2012.659463

Kloosterman, P. H., Keefer, K. V., Kelley, E. A., Summerfeldt, L. J., & Parker, J. D. (2011). Evaluation of the factor structure of the Autism-Spectrum Quotient. *Personality and Individual Differences, 50*(2), 310-314. https://doi.org/10.1016/j.paid.2010.10.015

Klusek, J., Martin, G. E., & Losh, M. (2014). A comparison of pragmatic language in boys with autism and fragile X syndrome. *Journal of Speech, Language, and Hearing Research*, *57*(5), 1692-1707. https://doi.org/10.1044/2014_JSLHR-L-13-0064

Kozak, M. J., & Cuthbert, B. N. (2016). The NIMH Research Domain Criteria Initiative: Background, Issues, and Pragmatics. *Psychophysiology, 53*(3), 286–297. https://doi.org/10.1111/psyp.12518

Kunihira, Y., Senju, A., Dairoku, H., Wakabayashi, A., & Hasegawa, T. (2006). 'Autistic' traits in non-autistic Japanese populations: relationships with personality traits and cognitive ability. *Journal of autism and developmental disorders*, *36*(4), 553-566. https://doi.org/10.1007/s10803-006-0094-1

Kurita, H., Koyama, T., & Osada, H. (2005). Autism-Spectrum Quotient–Japanese version and its short forms for screening normally intelligent persons with pervasive developmental disorders. *Psychiatry and clinical neurosciences, 59*(4), 490-496. https://doi.org/10.1111/j.1440-1819.2005.01403.x

L

Lam, K. S., Bodfish, J. W., & Piven, J. (2008). Evidence for three subtypes of repetitive behavior in autism that differ in familiality and association with other symptoms. *Journal of Child Psychology and Psychiatry, 49*(11), 1193-1200. https://doi.org/10.1111/j.1469-7610.2008.01944.x

Landa, R., Piven, J., Wzorek, M. M., Gayle, J. O., Chase, G. A., & Folstein, S. E. (1992). Social language use in parents of autistic individuals. *Psychological medicine*, *22*, 245-254. https://doi.org/10.1017/S0033291700032918

Landry, O., & Chouinard, P. A. (2016). Why we should study the broader autism phenotype in typically developing populations. *Journal of Cognition and*

Development, *17*(4), 584-595. https://doi.org/10.1080/15248372.2016.1200046

Le Couteur A, Bailey A, Goode S, Pickes A, Robertson S, Gottesman I, Rutter M. 1996. A broader phenotype of autism: The clinical spectrum in twins. *J Child Psychol Psychiat, 7*, 785–801. https://doi.org/10.1111/j.1469-7610.1996.tb01475.x

Lecavalier, L. (2005). An evaluation of the Gilliam autism rating scale. *Journal of autism and developmental disorders, 35*(6), 795-805. https://doi.org/10.1007/s10803-005-0025-6

Lepage, J. F., Lortie, M., Taschereau-Dumouchel, V., & Théoret, H. (2009). Validation of French-Canadian versions of the Empathy Quotient and Autism Spectrum Quotient. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 41*(4), 272. https://doi.org/10.1037/a0016248

Lever, A. G., Werkle-Bergner, M., Brandmaier, A. M., Ridderinkhof, K. R., & Geurts, H. M. (2015). Atypical working memory decline across the adult lifespan in autism spectrum disorder? *Journal of Abnormal Psychology, 124*(4), 1014–1026. https://doi.org/10.1037/abn0000108

Lin, J., Fu, X., Dai, S., Zeng, Y., Yang, R., Liu, R., ... & Ou, J. (2021). The Broad Autism Phenotype Questionnaire (BAPQ): Strengths, weaknesses and future improvements in Chinese version. *Asian Journal of Psychiatry, 60*, 102655. https://doi.org/10.1016/j.ajp.2021.102655

Lindgren, K. A., Folstein, S. E., Tomblin, J. B., & Tager-Flusberg, H. (2009). Language and reading abilities of children with autism spectrum disorders and specific language impairment and their first-degree relatives. *Autism Research*, *2*, 22-38. https://doi.org/10.1002/aur.63

Lombardo, M. V., Lai, M. C., Auyeung, B., Holt, R. J., Allison, C., Smith, P., ... & Wilson, C. E. (2016). Unsupervised data-driven stratification of mentalizing heterogeneity in autism. *Scientific Reports*, *6*, 1-15. https://doi.org/10.1038/srep35333

Lord C, Luyster RJ, Gotham K, Guthrie W. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part II): Toddler Module*. Torrance, CA: Western Psychological Services.

Lord C, Rutter M, DiLavore PC, Risi S, Gotham, K, & Bishop S. (2012). *Autism Diagnostic Observation Schedule, Second Edition*. Torrance, CA: Western Psychological Services.

Lord C, Rutter M, Le Couteur A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord, 24*(5), 659-85. https://doi.org/0162-325719411000--0659507.0

Losh, M., Adolphs, R., Poe, M. D., Couture, S., Penn, D., Baranek, G. T., & Piven, J. (2009). Neuropsychological profile of autism and the broad autism phenotype. *Archives of General Psychiatry, 66*, 518–526. https://doi.org/10.1001/archgenpsychiatry.2009.34

Losh, M., Childress, D., Lam, K., & Piven, J. (2008). Defining key features of the broad autism phenotype: A comparison across parents of multiple-and single-incidence autism families. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 147,* 424–433. https://doi.org/10.1002/ajmg.b.30612

Losh, M., Esserman, D., & Piven, J. (2010). Rapid automatized naming as an index of genetic liability to autism. *Journal of Neurodevelopmental Disorders*, *2*(2), 109- https://doi.org/116. 10.1007/s11689-010-9045-4

Losh, M., & Piven, J. (2007). Social-cognition and the broad autism phenotype: Identifying genetically meaningful phenotypes. *Journal of Child Psychology and Psychiatry*, *48*, 105-112. https://doi.org/10.1111/j.1469-7610.2006.01594.x

Lung, F. W., Shu, B. C., Chiang, T. L., & Lin, S. J. (2010). Parental concerns based general developmental screening tool and autism risk: the Taiwan National Birth cohort study. *Pediatric research*, *67*(2), 226-231. https://doi.org/10.1203/PDR.0b013e3181c8ed3d

Lyall, K., Schmidt, R. J., & Hertz-Picciotto, I. (2014). Maternal lifestyle and environmental risk factors for autism spectrum disorders. *International journal of epidemiology*, *43*(2), 443-464. https://doi.org/10.1093/ije/dyt282

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017

M

MacCallum, R. (1986). Specification searches in covariance structure modelling. *Psychological bulletin, 100*, 107-120. https://doi.org/10.1037/0033-2909.100.1.107

MacCallum, Robert C.; Roznowski, Mary; Necowitz, Lawrence B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111*(3), 490–504. https://doi.org/10.1037/0033-2909.111.3.490

Maes, J. H., Vissers, C. T., Egger, J. I., & Eling, P. A. (2013). On the relationship between autistic traits and executive functioning in a non-clinical Dutch student population. *Autism*, *17*(4), 379-389. https://doi.org/10.1177/1362361312442009

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7, 19–40. https://doi.org/10.1037/1082-989X.7.1.19

Mandell DS, Novak M. (2005) The role of culture in families' treatment decisions for children with autism spectrum disorders. *Ment Retard Dev Disabil Res Rev, 11*, 110–5. http://dx.doi.org/10.1002/mrdd.20061

Marquand, A. F., Wolfers, T., Mennes, M., Buitelaar, J., & Beckmann, C. F. (2016). Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. *Biological psychiatry: cognitive neuroscience and neuroimaging*, *1*(5), 433-447.

Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017). An overview of autism spectrum disorder, heterogeneity and treatment options. *Neuroscience bulletin*, *33*(2), 183-193.

Masters, G. N., & Wright, B. D. (1984). The essential process in a family of measurement models. *Psychometrika*, *49*(4), 529–544. https://doi.org/10.1007/bf02302590

Matheis, M., Matson, J. L., Burns, C. O., Jiang, X., Peters, W. J., Moore, M., ... & Estabillo, J. (2017). Factors related to parental age of first concern in toddlers with autism spectrum disorder. *Developmental Neurorehabilitation*, *20*(4), 228-235. https://doi.org/10.1080/17518423.2016.1211186

Matson, J. L., Matheis, M., Burns, C. O., Esposito, G., Venuti, P., Pisula, E., ... & Goldin, R. L. (2017). Examining cross-cultural differences in autism spectrum disorder: a multinational comparison from Greece, Italy, Japan, Poland, and the United States. *European Psychiatry*, *42*, 70-76. https://doi.org/10.1016/j.eurpsy.2016.10.007

Matson, J. L., & Williams, L. W. (2013). Differential diagnosis and comorbidity: distinguishing autism from other mental health issues. *Neuropsychiatry*, *3*(2), 233. https://doi.org/10.2217/npy.13.1

Matson, J. L., Worley, J. A., Fodstad, J. C., Chung, K. M., Suh, D., Jhin, H. K., ... & Furniss, F. (2011). A multinational study examining the cross-cultural differences in reported symptoms of autism spectrum disorders: Israel, South Korea, the United Kingdom, and the United States of America. *Research in Autism Spectrum Disorders*, *5*(4), 1598-1604. http://dx.doi.org/10.1016/j.rasd.2011.03.007

Meera, S. S., Girimaji, S. C., Seshadri, S. P., Philip, M., Shivashankar, N., Morgan, P., & Piven, J. (2015). Translation of the Broad Autism Phenotype Questionnaire to an Indian language: A description of the process. *Asian Journal of Psychiatry, 15*, 62–67. https://doi.org/10.1016/j.ajp.2015.04.013

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525-543. https://doi.org/10.1007/BF02294825

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist, 50*(9), 741. https://doi.org/10.1037/0003-066X.50.9.741

Micali, N., Hilton, K., Natatani, E., Heyman, I., Turner, C., & Mataix-Cols, D. (2011). Is childhood OCD a risk factor for eating disorders later in life? A longitudinal study. *Psychological Medicine*, *41*(12), 2507-2513. https://doi.org/10.1017/S003329171100078X

Milne, E., Gomez, R., Giannadou, A., & Jones, M. (2019). Atypical EEG in autism spectrum disorder: Comparing a dimensional and a categorical approach. *Journal of abnormal psychology*, *128*(5), 442. https://doi.org/10.1037/abn0000436

Miranda-Linné, F. M., & Melin, L. (2002). A factor analytic study of the Autism Behavior Checklist. *Journal of Autism and Developmental Disorders*, *32*(3), 181-188. https://doi.org/10.1023/A:1015519413133

Mitchell, M., & Jolley, J. (2013). *Research design explained* (8th ed.). Belmont, CA: Wadsworth.

Montiel-Nava, C., & Peña, J. A. (2008). Epidemiological findings of pervasive developmental disorders in a Venezuelan study. *Autism*, *12*(2), 191-202. https://doi.org/10.1177/1362361307086663

Morgan, B., Maybery, M., & Durkin, K. (2003). Weak central coherence, poor joint attention, and low verbal ability: Independent deficits in early autism. *Developmental psychology*, *39*(4), 646. https://doi.org/10.1037/0012-1649.39.4.646

Morrison, K. E., Chambers, L. K., Faso, D. J., & Sasson, N. J. (2018). The content and function of interests in the broad autism phenotype. *Research in Autism Spectrum Disorders*, *49*, 25-33. http://dx.doi.org/10.1016/j.rasd.2018.02.002

Moruzzi, S., Ogliari, A., Ronald, A., Happé, F., & Battaglia, M. (2011). The nature of covariation between autistic traits and clumsiness: a twin study in a general population sample. *Journal of autism and developmental disorders*, *41*(12), 1665-1674. http://dx.doi.org/10.1007/s10803-011-1199-8

Murphy, M., Bolton, P. F., Pickles, A., Fombonne, E., Piven, J., & Rutter, M. (2000). Personality traits of the relatives of autistic probands. *Psychological Medicine, 30*(6), 1411-1424. http://dx.doi.org/10.1017/S0033291799002949

N

Narayan, S., Moyes, B., & Wolff, S. (1990). Family characteristics of autistic children: A further report. *Journal of Autism and Developmental Disorders, 20*(4), 523-535. https://doi.org/10.1007/BF02216057

Nayar, K., Sealock, J. M., Maltman, N., Bush, L., Cook, E. H., Davis, L. K., & Losh, M. (2021). Elevated polygenic burden for autism spectrum disorder is associated with the broad autism phenotype in mothers of individuals with autism spectrum disorder. *Biological Psychiatry*, *89*(5), 476-485. https://doi.org/10.1016/j.biopsych.2020.08.029

Nguyen, P. H., Ocansey, M. E., Miller, M., Le, D. T., Schmidt, R. J., & Prado, E. L. (2019). The reliability and validity of the social responsiveness scale to measure autism symptomology in Vietnamese children. *Autism Research, 12*(11), 1706-1718. https://doi.org/10.1002/aur.2179

Nishiyama, T., Suzuki, M., Adachi, K., Sumi, S., Okada, K., Kishino, H., ... & Kanne, S. M. (2014). Comprehensive comparison of self-administered questionnaires for measuring quantitative autistic traits in adults. *Journal of autism and developmental disorders*, *44*(5), 993-1007. https://doi.org/10.1007/s10803-013-2020-7

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

O

Oliveira, G., Ataíde, A., Marques, C., Miguel, T. S., Coutinho, A. M., Mota-Vieira, L., ... & Vicente, A. M. (2007). Epidemiology of autism spectrum disorder in Portugal: prevalence, clinical characterization, and medical conditions. *Developmental Medicine & Child Neurology*, *49*(10), 726-733.https://doi.org/10.1111/j.1469-8749.2007.00726.x

Oreg, S. (2003). Resistance to change: Developing an individual differences measure. *Journal of applied psychology*, *88*(4), 680. https://doi.org/10.1037/0021-9010.88.4.680

O'Regan, D., & Tobiansky, R. (2014). Diagnosing autism spectrum disorders in older adults. *J Geriatr Med*, *10,* 46-51.

Osterlind, S. J. (1989). Style, Editorial, and Publication Guidelines for Items in the Multiple-Choice Format. In S. J. Osterlind (Ed). *Constructing Test Items* (pp. 173-217). Springer, Netherlands. https://doi.org/10.1007/0-306-47535-9

Ozonoff, S., Rogers, S. J., Farnham, J. M., & Pennington, B. F. (1993). Can standard measures identify subclinical markers of autism? *Journal of autism and developmental disorders*, *23*(3), 429-441.https://doi.org/10.1007/BF01046049

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., ... & Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: a Baby

Siblings Research Consortium study. *Pediatrics*, *128*(3), e488-e495. https://doi.org/10.1542/peds.2010-2825

P

Palmer, C. J., Paton, B., Enticott, P. G., & Hohwy, J. (2015). 'Subtypes' in the presentation of autistic traits in the general adult population. *Journal of Autism and Developmental Disorders, 45*(5), 1291-1301. https://doi.org/

Parkington, K. B., Clements, R. J., Landry, O., & Chouinard, P. A. (2015). Visual-motor association learning in undergraduate students as a function of the autism-spectrum quotient. *Experimental Brain Research, 233*, 2883– 2895. https://doi.org/10.1007/s00221-015-4358-x

Patra, S. (2016). Autism spectrum disorder in the elderly: a review of healthcare issues and challenges. *Journal of Geriatric Care and Research, 3*, 3-6.

Pérez-Albéniz, A., De Paúl, J., Etxeberría, J., Montes, M. P., & Torres, E. (2003). Adaptación de interpersonal reactivity index (IRI) al español. *Psicothema*, 267-272. Recuperado a partir de https://reunido.uniovi.es/index.php/PST/article/view/8118

Pisula, E., Kawa, R., Danielewicz, D., & Pisula, W. (2015). The relationship between temperament and autistic traits in a non-clinical student's sample. *PLoS One, 10*, e0124364. https://doi.org/10.1371/journal.pone.0124364

Piven, J. (2001). The broad autism phenotype: a complementary strategy for molecular genetic studies of autism. *American journal of medical genetics*, *105*, 34-35. https://doi.org/10.1002/1096-8628(20010108)105:1<34::AID-AJMG1052>3.0.CO;2-D

Piven, J., & Palmer, P. (1999). Psychiatric disorder and the broad autism phenotype: evidence from a family study of multiple-incidence autism families. American *Journal of Psychiatry, 156*(4), 557-563. https://doi.org/10.1176/ajp.156.4.557

Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997a). Personality and language characteristics in parents from multiple-incidence autism families. *American Journal of Medical Genetics, 74*(4), 398-411. https://doi.org/10.1002/(SICI)1096-8628(19970725)74:4<398::AID-AJMG11>3.0.CO;2-D

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997b). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry, 154*(2), 185–190. http://dx.doi.org/10.1176/ajp.154.2.185

Piven, J., Wzorek, M., Landa, R., Lainhart, J., Bolton, P., Chase, G. A., & Folstein, S. (1994). Personality characteristics of the parents of autistic individuals. *Psychological Medicine, 24*, 783–795. https://doi.org/10.1017/S0033291700027938.

Powell, P. S., Klinger, L. G., & Klinger, M. R. (2017). Patterns of age-related cognitive differences in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *47*(10), 3204-3219. https://doi.org/10.1007/s10803-017-3238-6

Preacher, K. J., MacCallum, R. C., Rucker, D. D., & Nicewander, W. A. (2005). Use of the extreme groups approach: A critical reexamination and new recommendations. *Psychological Methods, 10*(2), 178–192. https://doi.org/10.1037/1082-989X.10.2.178

Pruitt, M. M., Rhoden, M., & Ekas, N. V. (2018). Relationship between the broad autism phenotype, social relationships and mental health for mothers of children with autism spectrum disorder. *Autism*, *22*(2), 171-180. https://doi.org/10.1177/1362361316669621

R

Ratto, A. B., Reznick, J. S., Turner-Brown, L. (2015). Cultural effects on the diagnosis of autism spectrum disorder among Latinos. *Focus Autism Dev Disabil* http://dx.doi.org/10.1177/1088357615587501.1088357615587501.

Ridley, N. J., Homewood, J., & Walters, J. (2011). Cerebellar dysfunction, cognitive flexibility and autistic traits in a non-clinical sample. *Autism*, *15*(6), 728-745. https://doi.org/10.1177/1362361310395956

Robertson, A. E., & Simmons, D. R. (2013). The relationship between sensory sensitivity and autistic traits in the general population. *Journal of Autism and Developmental Disorders, 43*, 775–784. https://doi.org/10.1007/s10803-012-1608-7

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happe, F., et al. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Archives of General Psychiatry*, *68*(11), 1113-1121. https://doi.org/10.1001/archgenpsychiatry.2011.119.

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happé, F., ... & Ronald, A. (2012). A multivariate twin study of autistic traits in 12-year-olds: testing the fractionable autism triad hypothesis. *Behavior genetics*, *42*(2), 245-255. https://doi.org/10.1007/s10519-011-9500-3

Muñoz, F. J. R. (2012). La conciencia pragmática de adultos con síndrome de Asperger. *Revista de Logopedia, Foniatría y Audiología*, *32*, 21-31. https://doi.org/10.1016/j.rlfa.2012.01.002

Rogler LH. (1993). Culturally sensitizing psychiatric diagnosis. A framework for research. *J Nerv Ment Dis, 181*(7), 401–8. https://doi.org/10.1097/00005053-199307000-00001

Ronald, A., Happé, F., Bolton, P., Butcher, L. M., Price, T. S., Wheelwright, S., ... & Plomin, R. (2006). Genetic heterogeneity between the three components of the autism spectrum: a twin study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(6), 691-699. https://doi.org/10.1097/01.chi.0000215325.13058.9d

Ronald, A., Happe, F., Price, T. S., Baron-Cohen, S., & Plomin, R. (2006). Phenotypic and genetic overlap between autistic traits at the extremes of the general population. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(10), 1206-1214. https://doi.org/10.1097/01.chi.0000230165.54117.41

Ronald, A., & Hoekstra, R. (2014). Progress in understanding the causes of autism spectrum disorders and autistic traits: twin studies from 1977 to the present day. In Rhee S., Ronald A. (Eds). *Behavior genetics of psychopathology*. Advances in Behavior Genetics, vol. 2. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-9509-3_10

Ronald, A., Larsson, H., Anckarsäter, H., & Lichtenstein, P. (2011). A twin study of autism symptoms in Sweden. *Molecular Psychiatry*, *16*(10), 1039-1047. https://doi.org/10.1038/mp.2010.82

Ronemus, M., Iossifov, I., Levy, D., & Wigler, M. (2014). The role of de novo mutations in the genetics of autism spectrum disorders. *Nature Reviews Genetics*, *15*(2), 133-141. https://doi.org/10.1038/nrg3585

Rosenberg, R. E., Law, J. K., Yenokyan, G., McGready, J., Kaufmann, W. E., & Law, P. A. (2009). Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Archives of pediatrics & adolescent medicine, 163*(10), 907-914. https://doi.org/907-914. 10.1001/archpediatrics.2009.98

Rubenstein, E., & Chawla, D. (2018). Broader autism phenotype in parents of children with autism: a systematic review of percentage estimates. *Journal of child and family studies*, *27*(6), 1705-1720. https://doi.org/10.1007/s10826-018-1026-3

Russell, R. L., & Grizzle, K. L. (2008). Assessing child and adolescent pragmatic language competencies: Toward evidence-based assessments. *Clinical Child and Family Psychology Review*, *11*(1-2), 59-73. https://doi.org/10.1007/s10567-008-0032-1

Russell-Smith, S. N., Maybery, M. T., & Bayliss, D. M. (2011). Relationships between autistic-like and schizotypy traits: An analysis using the Autism Spectrum Quotient and Oxford-Liverpool Inventory of Feelings and Experiences. *Personality and Individual Differences, 51*(2), 128–132. https://doi.org/10.1016/j.paid.2011.03.027

Ruta, L., Mazzone, D., Mazzone, L., Wheelwright, S., & Baron-Cohen, S. (2012). The Autism-Spectrum Quotient—Italian version: A cross-cultural confirmation of the broader autism phenotype. *Journal of autism and developmental disorders, 42*(4), 625-633. https://doi.org/10.1007/s10803-011-1290-1

Rutter, M. (1978). Diagnosis and definition of childhood autism. *Journal of autism and childhood schizophrenia, 8*(2), 139-161. https://doi.org/10.1007/BF01537863

Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism Diagnostic Interview-Revised (ADI-R).* Los Angeles, CA: Western Psychological Services.

Ruzich, E., Allison, C., Smith, P., Ring, H., Auyeung, B., & Baron-Cohen, S. (2017). The autism-spectrum quotient in siblings of people with autism. Autism Research, 10(2), 289-297. https://doi.org/10.1002/aur.1651

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism, 6*, 2. https://doi.org/10.1186/2040-2392-6-2

S

Saito, Y., Suga, M., Tochigi, M., Abe, O., Yahata, N., Kawakubo, Y., ... & Yamasue, H. (2014). Neural correlate of autistic-like traits and a common allele in the oxytocin receptor gene. Social cognitive and affective neuroscience, 9(10), 1443-1450. https://doi.org/10.1093/scan/nst136

Sartawi, A. M. (1999). Educational and behavioural characteristics of autistic children in the United Arab Emirates. *International Journal of Rehabilitation Research*, *22*(4), 337-340.

Sasson, N. J., Faso, D. J., Parlier, M., Daniels, J. L., & Piven, J. (2014). When father doesn't know best: Selective disagreement between Self-Report and informant report of the broad autism phenotype in parents of a child with autism. *Autism Research*, *7*(6), 731-739. https://doi.org/10.1002/aur.1425

Sasson, N. J., Lam, K. S., Childress, D., Parlier, M., Daniels, J. L., & Piven, J. (2013a). The Broad Autism Phenotype Questionnaire: Prevalence and Diagnostic Classification. *Autism Research, 6*(2), 134-143. https://doi.org/10.1002/aur.1272

Sasson, N. J., Lam, K. S., Parlier, M., Daniels, J. L., & Piven, J. (2013b). Autism and the broad autism phenotype: familial patterns and intergenerational transmission. *Journal of Neurodevelopmental Disorders, 5,* 1-7. http://dx.doi.org/10.1186/1866-1955-5-11

Sasson, N. J., Nowlin, R. B., & Pinkham, A. E. (2013c). Social cognition, social skill, and the broad autism phenotype. *Autism, 17*, 655–667. https://doi.org/10.1177/1362361312455704

Scheeren, A. M., & Stauder, J. E. (2008). Broader autism phenotype in parents of autistic children: reality or myth?. *Journal of autism and developmental disorders*, *38*(2), 276. http://dx.doi.org/10.1007/s10803-007-0389-x

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002). The CAST (Childhood Asperger Syndrome Test) Preliminary development of a UK screen for mainstream primary-school-age children. *Autism*, *6*, 9-31. https://doi.org/10.1177/1362361302006001003

Seidman, I., Yirmiya, N., Milshtein, S., Ebstein, R. P., & Levi, S. (2012). The Broad Autism Phenotype Questionnaire: mothers versus fathers of children with an autism spectrum disorder. *Journal of autism and developmental disorders, 42*(5), 837-846. http://dx.doi.org/10.1007/s10803-011-1315-9

Shi, L.-J., Ou, J.-J., Gong, J.-B., Wang, S.-H., Zhou, Y.-Y., Zhu, F.-R., ... Luo, X.-R. (2015). Broad autism phenotype features of Chinese parents with autistic children and their associations with severity of social impairment in probands. *BMC Psychiatry*, *15*. https://doi.org/10.1186/s12888-015-0568-9

Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(8), 921-929. https://doi.org/10.1097/CHI.0b013e318179964f

Sipes, M., Furniss, F., Matson, J. L., Hattier, M. (2012). A multinational study examining the cross-cultural differences in social skills of children with autism spectrum disorders: a comparison between the United Kingdom and the United States of America. *J Dev Phys Disabil, 24,* 145–54. http://dx.doi.org/10.1007/ s10882-011-9261-1

Sireci, S. G. (1998). The construct of content validity. *Social indicators research, 45*, 83-117. https://doi.org/10.1023/A:1006985528729

Siu, W. K., Lam, C. W., Mak, C. M., Lau, E. T. K., Tang, M. H. Y., Tang, W. F., ... & Chan, K. Y. (2016). Diagnostic yield of array CGH in patients with autism spectrum disorder in Hong Kong. *Clinical and translational medicine*, *5*, 1-8. https://doi.org/10.1186/s40169-016-0098-1

Schaevitz, L. R., & Berger-Sweeney, J. E. (2012). Gene–environment interactions and epigenetic pathways in autism: the importance of one-carbon metabolism. ILAR journal, 53(3-4), 322-340. https://doi.org/10.1093/ilar.53.3-4.322

Skuse, D. H., Mandy, W. P., & Scourfield, J. (2005). Measuring autistic traits: heritability, reliability and validity of the Social and Communication

Disorders Checklist. *The British Journal of Psychiatry*, *187*(6), 568-572. https://doi.org/10.1192/bjp.187.6.568

Sterponi, L., & Shankey, J. (2014). Rethinking echolalia: Repetition as interactional resource in the communication of a child with autism. *Journal of child language, 41*(2), 275-304. https://doi.org/10.1017/S0305000912000682

Stewart, M. E., & Austin, E. J. (2009). The structure of the Autism-Spectrum Quotient (AQ): Evidence from a student sample in Scotland. *Personality and Individual Differences, 47*, 224-228. http://dx.doi.org/10.1016/j.paid.2009.03.004

Stewart, G. R., Charlton, R. A., & Wallace, G. L. (2018). Aging with elevated autistic traits: Cognitive functioning among older adults with the broad autism phenotype. *Research in Autism Spectrum Disorders*, *54*, 27-36. https://doi.org/10.1016/j.rasd.2018.06.009

Stewart, M. E., & Ota, M. (2008). Lexical effects on speech perception in individuals with "autistic" traits. *Cognition*, *109*, 157-162. https://doi.org/10.1016/j.cognition.2008.07.010

Stewart, M. E., Watson, J., Allcock, A. J., & Yaqoob, T. (2009). Autistic traits predict performance on the block design. *Autism*, *13*(2), 133-142. https://doi.org/10.1177/1362361308098515

Stojković, I., Dučić, B., Kaljača, S., & Đorđević, M. (2018). Construction and psychometric evaluation of a short form of the Broad Autism Phenotype Questionnaire. *Psihologija, 51*(2), 243-258. https://doi.org/10.2298/PSI170619014S

Sucksmith, E., Roth, I., & Hoekstra, R. A. (2011). Autistic traits below the clinical threshold: re-examining the broader autism phenotype in the 21st century. *Neuropsychology Review*, *21*(4), 360–389. https://doi.org/10.1007/s11065-011-9183-9.

Suen, H., & Ary, D. (1989). *Analyzing Quantitative Behavioral Observation Data*. Psychology Press. https://doi.org/10.4324/9781315801827

Sun, X., Allison, C., Auyeung, B., Matthews, F. E., Zhang, Z., Baron-Cohen, S., & Brayne, C. (2014). Comparison between a mandarin Chinese version of the childhood autism spectrum test and the Clancy autism behaviour scale in mainland

China. *Research in developmental disabilities, 35*(7), 1599-1608. https://doi.org/10.1016/j.ridd.2014.02.005

Sung, Y. J., Dawson, G., Munson, J., Estes, A., Schellenberg, G. D., & Wijsman, E. M. (2005). Genetic investigation of quantitative traits related to autism: use of multivariate polygenic models with ascertainment adjustment. *The American Journal of Human Genetics*, *76*, 68-81. https://doi.org/10.1086/426951

Szatmari, P., Georgiades, S., Bryson, S., Zwaigenbaum, L., Roberts, W., Mahoney, W., ... & Tuff, L. (2006). Investigating the structure of the restricted, repetitive behaviours and interests domain of autism. *Journal of Child Psychology and Psychiatry, 47*(6), 582-590. https://doi.org/10.1111/j.1469-7610.2005.01537.x

Szatmari, P., Georgiades, S., Duku, E., Zwaigenbaum, L., Goldberg, J., & Bennett, T. (2008). Alexithymia in parents of children with autism spectrum disorder. *Journal of autism and developmental disorders*, *38*(10), 1859-1865.https://doi.org/10.1007/s10803-008-0576-4

Szatmari, P., Jones, M. B., Fisman, S., Tuff, L., Bartolucci, G., Mahoney, W., & Bryson, S. E. (1995). Parents and collateral relatives of children with pervasive developmental disorders: A family history study. *American Journal of Medical Genetics, 60,* 282–289. https://doi.org/10.1002/ajmg.1320600405

Szatmari, P., Paterson, A. D., Zwaigenbaum, L., Roberts, W., Brian, J., Liu, X. Q., ... & Autism Genome Project Consortium. (2007). Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nature genetics*, *39*(3), 319.https://doi.org/10.1038/ng1985

Szmigiera, M. (2021, 30 marzo). The most spoken languages worldwide 2021. *Statista 2021*. https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/

T

Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B. R. I. A. N., Putnam, S. A. R. A., McGrath, L., ... & Folstein, S. E. (2003). A principal components analysis of the Autism Diagnostic Interview-Revised. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*(7), 864-872. https://doi.org/10.1097/01.CHI.0000046870.56865.90

Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. (2014). Evidence for distinct cognitive profiles in autism spectrum disorders and specific language impairment. *Journal of Autism and Developmental Disorders, 44*, 19-30. https://doi.org/10.1007/s10803-013-1847-2

Tilden, V. P., Nelson, C. A., & May, B. A. (1990). Use of qualitative methods to enhance content validity. *Nursing Research*. https://doi.org/10.1097/00006199-199005000-00015

Tsang, T., Gillespie-Lynch, K., & Hutman, T. (2016). Theory of mind indexes the broader autism phenotype in siblings of children with autism at school age. *Autism Research and Treatment,* 1-13. https://doi.org/10.1155/2016/6309189.

V

Valla, J. M., Ganzel, B. L., Yoder, K. J., Chen, G. M., Lyman, L. T., Sidari, A. P., ... & Belmonte, M. K. (2010). More than maths and mindreading: sex differences in empathizing/systemizing covariance. *Autism Research, 3*(4), 174-184. https://doi.org/10.1002/aur.143

Van Lang, N. D., Boomsma, A., Sytema, S., de Bildt, A. A., Kraijer, D. W., Ketelaars, C., & Minderaa, R. B. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *Journal of Child Psychology and Psychiatry, 47*, 37-44. https://doi.org/10.1111/j.1469-7610.2005.01434.x

Van Niekerk, M. E., Groen, W., Vissers, C. T. W., van Driel-de Jong, D., Kan, C. C., & Voshaar, R. C. O. (2011). Diagnosing autism spectrum disorders in elderly people. *International Psychogeriatrics, 23*(5), 700-710. https://doi.org/10.1017/S1041610210002152

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. Journal of Research in Personality, 40(5), 472-481. https://doi.org/10.1016/j.jrp.2005.03.003

Virkud, Y. V., Todd, R. D., Abbacchi, A. M., Zhang, Y., & Constantino, J. N. (2009). Familial aggregation of quantitative autistic traits in multiplex versus simplex autism. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 150*(3), 328-334. https://doi.org/10.1002/ajmg.b.30810

Volkmar, F. R., Paul, R., Rogers, S. J., & Pelphrey, K. A. (Eds.). (2014). *Handbook of autism and pervasive developmental disorders, diagnosis, development, and brain mechanisms* (2nd Edition). John Wiley & Sons.

W

Wainer, A. L., Block, N., Donnellan, M. B., & Ingersoll, B. (2013). The broader autism phenotype and friendships in non-clinical dyads. *Journal of autism and developmental disorders*, *43*(10), 2418-2425. https://doi.org/10.1007/s10803-013-1789-8

Wainer, A. L., Ingersoll, B. R., & Hopwood, C. J. (2011). The Structure and Nature of the Broader Autism Phenotype in a Non-Clinical Sample. *Journal of Psychopathology and Behavioral Assessment, 33*, 459–469. http://dx.doi.org/10.1007/s10862-011-9259-0

Wallace, G. L., Budgett, J., & Charlton, R. A. (2016). Aging and autism spectrum disorder: Evidence from the broad autism phenotype. *Autism Research*, *9*(12), 1294-1303. https://doi.org/10.1002/aur.1620

Wakabayashi, A., Baron-Cohen, S., & Wheelwright, S. (2006). Are autistic traits an independent personality dimension? A study of the Autism-Spectrum Quotient (AQ) and the NEO-PI-R. *Personality and individual differences*, *41*(5), 873-883. https://doi.org/10.1016/j.paid.2006.04.003

Wang, T., Guo, H., Xiong, B., Stessman, H. A., Wu, H., Coe, B. P., ... & Eichler, E. E. (2016). De novo genic mutations among a Chinese autism spectrum disorder cohort. *Nature communications*, *7*, 1-10. https://doi.org/10.1038/ncomms13316

Warrier, V., Chee, V., Smith, P., Chakrabarti, B., & Baron-Cohen, S. (2015). A comprehensive meta-analysis of common genetic variants in autism spectrum conditions. *Molecular autism*, *6*, 1-11. https://doi.org/10.1186/s13229-015-0041-0

Wheelwright, S., Auyeung, B., Allison, C., & Baron-Cohen, S. (2010). Defining the broader, medium, and narrow autism phenotype among parents using the Autism Spectrum Quotient (AQ). *Molecular Autism, 1,* 1–10. https://doi.org/10.1186/2040-2392-1-10

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., ... & Wakabayashi, A. (2006). Predicting autism spectrum quotient (AQ) from the systemizing quotient-revised (SQ-R) and empathy quotient (EQ). *Brain research*, *1079*, 47-56. https://doi.org/10.1016/j.brainres.2006.01.012

Whitehouse, A. J., Barry, J. G., & Bishop, D. V. (2007). The broader language phenotype of autism: a comparison with specific language impairment. *Journal of Child Psychology and Psychiatry*, *48*(8), 822-830. https://doi.org/10.1111/j.1469-7610.2007.01765.x

Whitehouse, A. J. O., Coon, H., Miller, J., Salisbury, B., & Bishop, D. V. M. (2010). Narrowing the broader autism phenotype: A study using the Communication Checklist—Adult Version (CC-A). *Autism, 14*(6), 559–574. https://doi.org/10.1177/1362361310382107

Williams, J., Scott, F., Stott, C., Allison, C., Bolton, P., Baron-Cohen, S., & Brayne, C. (2005). The CAST (childhood asperger syndrome test) test accuracy. *Autism, 9*, 45-68. https://doi.org/10.1177/1362361305049029

Wing L. (1988) The Continuum of Autistic Characteristics. In: Schopler E., Mesibov G.B. (Eds). *Diagnosis and Assessment in Autism. Current Issues in Autism.* Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-0792-9_7

Wiśniowiecka-Kowalnik, B., & Nowakowska, B. A. (2019). Genetics and epigenetics of autism spectrum disorder—current evidence in the field. *Journal of applied genetics*, *60*, 37-47. https://doi.org/10.1007/s13353-018-00480-w

Wood, J. J., Ehrenreich-May, J., Alessandri, M., Fujii, C., Renno, P., Laugeson, E., ... & Storch, E. A. (2015). Cognitive behavioral therapy for early adolescents with autism spectrum disorders and clinical anxiety: A randomized, controlled trial. *Behavior therapy*, *46*, 7-19. https://doi.org/10.1016/j.beth.2014.01.002

World Health Organization. (2019). *International Statistical Classification of Diseases and Related Health Problems (CID-11)*. World Health Organization.

World Health Organization. (2019). *International Statistical Classification of Diseases and Related Health Problems (CID-11)*. World Health Organization.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis.* MESA press.

Z

Zhang, L., Sun, Y., Chen, F., Wu, D., Tang, J., Han, X., ... & Wang, K. (2016). Psychometric properties of the Autism-Spectrum Quotient in both clinical and non-clinical samples: Chinese version for mainland China. *BMC psychiatry, 16*, 1-10. https://doi.org/10.1186/s12888-016-0915-5

APPENDICES

Appendix 1. Chapter Three. Study 1. Psychometric Properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements

BAPQ-SP QUESTIONNAIRE

Instructions and Items of the Original and Adapted Version

BAPQ-SP	**Original BAPQ (Hurley et al. 2007)**
Para cada pregunta, **marque con una (X)** el valor que mejor describe cuán a menudo esa frase se aplica a usted según los valores de la siguiente lista:	For each question, **mark with a (X)** the answer that best describes how often that statement applies to you on a scale of:
1. Muy raramente · **2. Raramente** · **3. Ocasionalmente** · **4. Bastante a menudo** · **5. A menudo** · **6. Muy a menudo**	**1. Very rarely** · **2. Rarely** · **3. Occasionally** · **4. Somewhat often** · **5. Often** · **6. Very often**
Las preguntas **marcadas con asterisco (*)** hacen referencia a interacciones casuales con conocidos (no con amigos íntimos o familiares).	The questions **marked with asterisk (*)** refer to casual interaction with acquaintances (no with close friends and family members).
Preguntas:	**Questions:**
1. Me gusta estar con otras personas.	1. I like being around other people.
2. Me resulta difícil expresarme con fluidez.	2. I find it hard to get my words out smoothly.
3. Me siento cómodo/a con los cambios inesperados de planes.	3. I am comfortable with unexpected changes in plans.
4. Me cuesta seguir el curso de las conversaciones.	4. It's hard for me to avoid getting side-tracked in conversation.
5. Prefiero hablar con la gente para pedir información antes que para socializarme.	5. I would rather talk to people to get information than to socialize.
6. La gente tiene que convencerme para que pruebe algo nuevo.	6. People have to talk me into trying something new.
7. Mientras mantenemos una conversación congenio con la otra persona *.	7. I am ''in-tune'' with the other person during conversation *.
8. Tengo que mentalizarme ante la idea de visitar un lugar que no me sea familiar.	8. I have to warm myself up to the idea of visiting an unfamiliar place.
9. Disfruto estando en situaciones sociales.	9. I enjoy being in social situations.
10. Cuando converso con alguien mi voz tiene una entonación plana o monótona.	10. My voice has a flat or monotone sound to it.

11. Al conversar con otras personas me siento fuera de lugar *.	11. I feel disconnected or "out of sync" in conversations with others *.
12. A la gente le parece fácil acercarse a mí *.	12. People find it easy to approach me *.
13. Tengo una fuerte necesidad de que mis días discurran de forma muy similar.	13. I feel a strong need for sameness from day to day.
14. La gente me pide que repita lo que he dicho porque no han entendido lo que he querido decir.	14. People ask me to repeat things I've said because they don't understand.
15. Soy flexible sobre cómo deberían hacerse las cosas.	15. I am flexible about how things should be done.
16. Me gusta buscar situaciones en las que pueda conocer gente nueva.	16. I look forward to situations where I can meet new people.
17. Me han dicho que hablo demasiado sobre ciertos temas.	17. I have been told that I talk too much about certain topics.
18. Cuando doy conversación es sólo para ser educado/a *.	18. When I make conversation it is just to be polite *.
19. Tengo ganas de probar cosas nuevas.	19. I look forward to trying new things.
20. Hablo demasiado fuerte o demasiado bajo.	20. I speak too loudly or softly.
21. Sé cuándo alguien no está interesado en lo que estoy diciendo *.	21. I can tell when someone is not interested in what I am saying *.
22. Tengo dificultades para enfrentarme a cambios en mi rutina.	22. I have a hard time dealing with changes in my routine.
23. Prefiero establecer interacciones superficiales *.	23. I am good at making small talk *.
24. Soy muy rígido/a en mis costumbres.	24. I act very set in my ways.
25. Siento que congenio realmente con otras personas.	25. I feel like I am really connecting with other people.
26. La gente se frustra debido a mi reticencia o falta de voluntad para ceder.	26. People get frustrated by my unwillingness to bend.
27. Dar conversación me aburre *.	27. Conversation bores me *.
28. Soy cariñoso/a y amable en mis interacciones con los demás.	28. I am warm and friendly in my interactions with others *.
29. Durante una conversación con otras personas suelo dejar largas pausas.	29. I leave long pauses in conversation.
30. Altero mi rutina diaria probando algo diferente.	30. I alter my daily routine by trying something different.
31. Prefiero estar solo/a en lugar de acompañado.	31. I prefer to be alone rather than with others.

32. Cuando hablo con la gente tengo tendencia a irme por las ramas.	32. I lose track of my original point when talking to people.
33. Me gusta seguir una rutina minuciosa cuando trabajo.	33. I like to closely follow a routine while working.
34. Cuando estoy conversando me doy cuenta de cuándo es el momento de cambiar de tema *.	34. I can tell when it is time to change topics in conversation *.
35. Sigo haciendo las cosas de la manera que sé, incluso aunque haya otro modo mejor de hacerlas.	35. I keep doing things the way I know, even if another way might be better.
36. Disfruto charlando con la gente *.	36. I enjoy chatting with people *.

Appendix 2. Chapter Five. Study 3

"Is It Possible to Assess the Two-domain Definition of the Broad Autism Phenotype Using the Available Measurement Tools?"

Table 1.

Proposed BAP Operationalization

Broad Autism Phenotype Definition (BAP)	
A. Social interaction and social communication impairments (SCI BAP) : Persistent deficits in social communication and social interaction across multiple contexts.	
A1. Deficits in social-emotional reciprocity	From abnormal social approach and failure of normal back-and-forth conversation, reduced sharing of interests, emotions, or affect to failure to initiate or respond to social interactions.
A2. Deficits in non-verbal communicative behaviours used for social interaction	From poorly integrated verbal and nonverbal communication to abnormalities in eye contact and body language, deficits in understanding and use of gestures to a total lack of facial expressions and nonverbal communication.
A3. Deficits in developing, maintaining, and understanding relationships	From difficulties adjusting behaviours to suit various social contexts to difficulties in sharing imaginative play or to difficulties in making friends or to the absence of interest in peers.
B. Restricted/Repetitive Behaviours, Interests, or Activities (RRB BAP): Restricted and/or repetitive patterns of behaviours, interests, or activities.	
B1. Stereotyped or repetitive motor movements, use of objects, or speech	Simple motor stereotypes, lining up toys or flipping objects, Echolalia, Idiosyncratic phrases.
B2. Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behaviour	Extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take the same route or eat the same food every day.
B3. Highly restricted, fixated interests	Hypersensitivity to criticism, excessively circumscribed or perseverative interests, strong attachment to or preoccupation with unusual objects.
B4. Hyper- or hyporeactivity to sensory input or unusual interest in sensory aspects of the environment	Apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement.

Table 2.

Items Selected in Each Phase

Subdomain	Test	Assignation of the items	Itemmetric analysis	Expert judgment
A. Social interaction and social communication impairments (SCI BAP): Persistent deficits in social communication and social interaction across multiple contexts.				
A1. Deficits in social-emotional reciprocity	BAPQ	2, 4, 7, 11, 12, 14, 25, 27, 28, 32, 36	2, 4, 7, 11, 14, 27, 32, 36	36
	SRS-2	3, 5, 7, 10, 12, 13, 19, 26, 34, 35, 38, 43, 48, 51, 60	3, 5, 7, 10, 12, 13, 26, 34, 35, 38, 43, 51, 60	5, 7, 10, 12, 13, 26, 35, 38, 60
	AQ	10, 17, 18, 26, 33, 35, 38	10, 17, 33, 38	10
A2. Deficits in non-verbal communicative behaviours used for social interaction	BAPQ	10, 20, 21, 29, 34	10, 21, 29, 34	10, 21, 34
	SRS-2	2, 15, 16, 46, 52, 53	2, 46, 52, 53, 27,	2, 53
	AQ	27, 31, 36, 45	31, 36, 45	27, 31, 36, 45
A3. Deficits in developing, maintaining, and understanding relationships	BAPQ	1, 5, 9, 16, 18, 23, 31	1, 18	1
	SRS-2	1, 6, 18, 21, 22, 23, 27, 33, 36, 37, 47, 55, 56, 57, 64	21, 22, 27, 36, 37, 47, 56, 64,	37, 47, 64,
	AQ	1, 7, 8, 11, 15, 20, 22, 40, 42, 44, 47, 48, 50	1, 8 ,20, 40, 42, 44, 47, 48, 50	1, 20, 40, 44, 47, 50
B. Restricted/Repetitive Behaviours, Interests, or Activities (RRB BAP): Restricted and/or repetitive patterns of behaviours, interests, or activities.				
Subdomain	**Test**	**Assignation of the items**	**Itemmetric analysis**	**Expert judgment**
B1. Stereotyped or repetitive motor movements, use of objects, or speech	BAPQ	0	0	0
	SRS-2	50	50	50
	AQ	0	0	0

B2. Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behaviour	BAPQ	3, 6, 8, 13, 15, 19, 22, 24, 26, 30, 33, 35	15, 22, 24, 33	15, 22, 24, 33
	SRS-2	4, 24, 61, 63	24, 61	24, 61
	AQ	2, 25, 34, 43, 46	2, 25, 43, 46	2, 25, 46
B3. Highly restricted, fixated interests	BAPQ	17	17	17
	SRS-2	28, 31, 39	28, 31	28, 31
	AQ	4, 6, 9, 16, 19, 23, 39, 41	19, 39, 41	39, 41
B4. Hyper- or hyporeactivity to sensory input or unusual interest in sensory aspects of the environment	BAPQ	0	0	0
	SRS-2	20, 42	20, 42	20, 42
	AQ	5, 12	5, 12	12

Table 3.

Items Discarded Due to Itemmetric Problems in Each Property

	Itemmetric properties				
Questionnaire	**Clarity**	**Comprehensibility**	**Concreteness**	**Degree of self-reference**	**Evaluation of the items**
BAPQ	13, 20, 25, 35	26	3, 5, 6, 8, 9, 12, 16, 19, 23, 28, 30, 31, 35	12, 26	5, 23
SRS-2	4, 11, 12, 13, 18, 30, 35, 40	45	1, 4, 6, 14, 30	0	1, 4, 12, 15, 24, 36, 42, 45
AQ	3, 6, 13, 15, 21, 25		5, 8, 10, 11, 12, 19, 23, 25, 28	15	6, 9, 12, 18

Appendix 3. Chapter Five. Study 3. Document 1

[This page is intentionally left blank]

CUESTIONARIO DE EXPERTOS/AS PARA LA CONSTRUCCIÓN DE UN TEST INTEGRADO DEL FENOTIPO AMPLIADO AUTISTA

***Guía y conceptos básicos*.**

Es recomendable tener esta guía al alcance a medida que se contesta el cuestionario, en ella vienen recogidas las instrucciones necesarias e información complementaria para dicha tarea.

Carta de presentación

Distinguido/a profesional:

Nos dirigimos a usted dada su amplia experiencia con personas diagnosticadas con Trastornos del Espectro Autista (TEA) y con familiares de los mismos. El objeto de esta comunicación es solicitar su colaboración, en calidad de experto/a, para la construcción del primer instrumento de medida de las expresiones del Fenotipo Ampliado del Autismo (FAA) en la población española. Para ello, el primer paso consiste en seleccionar de entre las herramientas más relevantes a nivel internacional, cuáles serían los ítems más adecuados para captar las dimensiones y subdimensiones del TEA definidas en el DSM-V, **pero aplicadas a población subclínica española.**

Antes de continuar, permítanos definirle el concepto de FAA; el FAA hace referencia a un conjunto de características **subclínicas** cualitativamente similares a aquellas presentadas por las personas diagnosticadas con Trastorno del Espectro Autista (TEA), pero más sutiles en su expresión. Inicialmente, el FAA fue observado en padres de niños ya diagnosticados con TEA y en familiares de primer orden (cercanos). Sin embargo, recientes estudios apoyan su existencia también en población general. En la actualidad, son tres los instrumentos utilizados en la literatura para medir FAA: el Autism Quotient (AQ; Baron-Cohen et al., 2001), el Social Responsiveness Scale (SRS; Constantino, 2007), y el Broad Autism Phenotype questionnaire (BAPQ; Hurley et al., 2007). Sin embargo, estas herramientas de medida o bien no se encuentran disponibles para su aplicación en la población española o bien no presentan adecuadas propiedades psicométricas.

Consecuentemente, en nuestro estudio nos proponemos crear una nueva herramienta para la medición del FAA en sintonía con la nueva conceptualización del DSM-V (APA, 2013) y adaptada a la población española, a partir de la selección de los ítems más adecuados de los cuestionarios ya mencionados. El **Objetivo General** para el que se empleará el test será el de diferenciar entre distintitos grados de severidad de FAA en cada una de las subdimensiones del TEA, tanto en familiares de niños diagnosticados con TEA como en población general.

Su ayuda es un pilar fundamental en esta fase del estudio. Solo una adecuada muestra de expertos/as en todas las áreas que tratan con personas diagnosticadas con TEA, o con características salientes del FAA, permitirá seleccionar de forma neutral y precisa cuáles son los ítems más adecuados para llevar a cabo una medición eficiente y fiable que permita la detección de los rasgos más sutiles del FAA.

Dada la importancia de su colaboración, además de mostrarle nuestro más sincero agradecimiento, nos gustaría ponernos a su disposición para cualquier tipo de consulta, colaboración o ayuda que podamos brindarle desde nuestro equipo de investigación. Adicionalmente, nos gustaría indicarle que, si así lo desea, le mantendremos informado sobre el desarrollo de la herramienta de medida y le facilitaremos dicha herramienta y su baremo una vez hayamos finalizado el proceso de construcción. Asimismo, como no podría ser menos, en caso de que finalmente decida colaborar con nuestro equipo, su nombre será incluido en los agradecimientos del artículo de investigación que describa el proceso de construcción de la herramienta de medida.

Atentamente,

M. Ángeles Estévez, investigadora principal del proyecto, perteneciente al grupo de investigación HUM891 (Investigación en Neurociencia Cognitiva) y profesora titular de la Universidad de Almería

Instrucciones para llevar a cabo el juicio de expertos

Para llevar a cabo el siguiente procedimiento es necesario que usted tenga claros los siguientes aspectos:

1. La **población** a la que está dirigida el test es tanto **población general** como **familiares de personas diagnosticadas con TEA.**
2. El **objetivo** último del test es poder **diferenciar**, a las personas que contesten al test, **en función de la severidad con la que expresan las características principales de cada _subdimensión_** (sin olvidar que estas expresiones se expresarán sutilmente o a nivel subclínico).
3. La estructura del test que se pretende desarrollar coincide con la definición del constructo del TEA en el DSM-V (APA, 2013): dos dimensiones principales con tres y cuatro subdimensiones respectivamente.

TABLA DE CONTENIDO	
Dimensión	**Subdimensión**
Problemas en la comunicación e interacción social	**1) Deficiencias en la reciprocidad socioemocional (es decir, problemas en la interacción y comunicación socioemocional).**
	2) Deficiencias en las conductas comunicativas no verbales utilizadas en la interacción social.
	3) Deficiencias en el desarrollo, mantenimiento y comprensión de las relaciones (es decir, saber cómo comportarse en diferentes situaciones sociales, dificultades para hacer amigos y ausencia de interés por otras personas).
Patrones de intereses obsesivos y conductas repetitivas	**1) Movimientos, utilización de objetos o habla estereotipados o repetitivos.**
	2) Insistencia en la monotonía, excesiva inflexibilidad de rutinas o patrones ritualizados de comportamiento verbal o no verbal.
	3) Intereses muy restringidos y fijos que son anormales en cuanto a su intensidad o foco de interés.
	4) Híper o hiporreactividad a los estímulos sensoriales o interés inhabitual por aspectos sensoriales del entorno.

4. La interpretación de las puntuaciones del test se llevará a cabo usando modelos de teoría de respuesta al ítem. No estamos buscando ítems que sean paralelos en cada subdimensión (que midan exactamente lo mismo). Pretendemos incluir ítems que expresen **distinto grado de severidad en cada *subdimensión*.**
5. Cada vez que emita un juicio o conteste a una pregunta referida al test tiene que **tener en mente la población a la que está dirigido** (tanto población general como familiares de personas diagnosticadas con TEA) **y el objetivo** último del test (diferenciar, a **nivel subclínico**, a las personas que contesten al test en función de la severidad con la que expresan cada ***subdimensión***).
6. Tendrá que realizar juicios sobre el test, sus instrucciones, los ítems y sus opciones de respuesta basados en los siguientes conceptos:
    a. **Relevancia**: importancia/relación con la subdimensión destacada. Esta característica indica en qué medida el ítem refleja comportamientos o indicadores destacados de cada subdimensión.
    b. **Adecuación a los objetivos**: grado en el que cumple las necesidades del estudio. Cada ítem pretende diferenciar a las personas que tengan distintos niveles de FAA **expresados específicamente en cada *subdimensión*** de la forma más eficiente posible (es importante recordar que la población destino será población general o familiares de niños/as diagnosticados con TEA, es decir, personas cuya mayor severidad en la expresión del TEA será a nivel subclínico). En los ítems poco adecuados todas las personas responderán casi de la misma manera (generalmente las categorías centrales), independientemente de sus niveles de FAA.
    c. **Respuesta diferencial en función de distintos grupos:** grado en el que el contenido del ítem puede generar respuestas diferentes SOLO debidas al grupo de pertenencia (o bien población general, o bien padres/familiares de niños/as diagnosticados con TEA) y no debidas a los distintos niveles de FAA de las personas que contestan.

Appendix 4. Chapter Five. Study 3. Document 2

[This page is intentionally left blank]

Questionario de expertos/as para la construcción de un test integrado del Fenotipo Ampliado Autista

Estudio sobre la construcción de una nueva herramienta de medida que evalúe FAA en población española.

[This page is intentionally left blank]

DIMENSIÓN 1: PROBLEMAS EN LA COMUNICACIÓN E INTERACCIÓN SOCIAL.
***SUBDIMENSIÓN 1*: DEFICIENCIAS EN LA REPCIPROCIDAD SOCIOEMOCIONAL**

Opciones de respuesta: *1 (ninguna), 2 (leve), 3 (moderada) y 4 (totalmente).*

Ítem		**Relevancia para la sub-dimensión**				**Adecuación a los objetivos**				**Respuesta diferencial en función de distintos grupos**			
1a	**Al conversar con otras personas me siento fuera de lugar.**	1	2	3	4	1	2	3	4	1	2	3	4
1b	**Disfruto estando de cháchara (hablando de trivialidades).**	1	2	3	4	1	2	3	4	1	2	3	4
1c	**Tengo dificultades para responder preguntas de manera directa y acabo yéndome por las ramas.**	1	2	3	4	1	2	3	4	1	2	3	4
1d	**Mientras mantenemos una conversación congenio con la otra persona.**	1	2	3	4	1	2	3	4	1	2	3	4
1e	**Disfruto de las conversaciones triviales (conversaciones informales con otras personas).**	1	2	3	4	1	2	3	4	1	2	3	4
1f	**Dar conversación me aburre.**	1	2	3	4	1	2	3	4	1	2	3	4
1g	**Disfruto charlando con la gente.**	1	2	3	4	1	2	3	4	1	2	3	4
1h	**Normalmente soy consciente de cómo se sienten los demás.**	1	2	3	4	1	2	3	4	1	2	3	4
1i	**Evito a la gente que intenta acercarse emocionalmente a mí.**	1	2	3	4	1	2	3	4	1	2	3	4
1j	**Me cuesta seguir el curso de las conversaciones.**	1	2	3	4	1	2	3	4	1	2	3	4
1k	**La gente me pide que repita lo que he dicho porque no han entendido lo que he querido decir.**	1	2	3	4	1	2	3	4	1	2	3	4

1l	**Me siento seguro/a de mí mismo/a cuando interactúo con otras personas.**	1	2	3	4	1	2	3	4	1	2	3	4
1m	**Cuando hablo con la gente tengo tendencia a irme por las ramas.**	1	2	3	4	1	2	3	4	1	2	3	4

DIMENSIÓN 1: PROBLEMAS EN LA COMUNICACIÓN E INTERACCIÓN SOCIAL.
***SUBDIMENSIÓN 1**: DEFICIENCIAS EN LA REPCIPROCIDAD SOCIOEMOCIONAL*

Opciones de respuesta: *1 (ninguna), 2 (leve), 3 (moderada) y 4 (totalmente).*

Ítem	Relevancia para la sub-dimensión				Adecuación a los objetivos				Respuesta diferencial en función de distintos grupos			
1n Respondo de manera apropiada a los cambios de humor de los demás (e.g., cuando un amigo pasa de estar contento a estar triste).	1	2	3	4	1	2	3	4	1	2	3	4
1o En situaciones sociales, tengo facilidad para seguir el hilo de diferentes conversaciones.	1	2	3	4	1	2	3	4	1	2	3	4
1p En una conversación telefónica, no estoy seguro/a de cuándo me toca hablar.	1	2	3	4	1	2	3	4	1	2	3	4
1q Tengo problemas para mantener el curso de una conversación.	1	2	3	4	1	2	3	4	1	2	3	4
1r Ofrezco consuelo a los demás cuando se sienten tristes.	1	2	3	4	1	2	3	4	1	2	3	4
1s Los demás piensan que soy emocionalmente distante y que no muestro mis sentimientos.	1	2	3	4	1	2	3	4	1	2	3	4
1t Se me da bien la cháchara (hablar de trivialidades).	1	2	3	4	1	2	3	4	1	2	3	4
1u Soy capaz de expresar mis sentimientos a otras personas.	1	2	3	4	1	2	3	4	1	2	3	4
1v Soy torpe con los turnos en las interacciones con los demás (e.g., me cuesta trabajo mantener. El "toma y daca de una conversación).	1	2	3	4	1	2	3	4	1	2	3	4
1w Me tomo las cosas al pie de la letra, y por eso malinterpreto el significado	1	2	3	4	1	2	3	4	1	2	3	4

	de algunas partes de la conversación.												
1x	**Me resulta difícil expresarme con fluidez.**	1	2	3	4	1	2	3	4	1	2	3	4
1y	**No me doy cuenta cuándo los demás están tratando de aprovecharse de mí.**	1	2	3	4	1	2	3	4	1	2	3	4

- **¿Considera que alguno de los anteriores ítems podría ser más adecuado para medir cualquier otra *subdimensión* distinta de esta?** Consulte la Tabla de contenido del apartado 3 de las Instrucciones para llevar a cabo el juicio de expertos y rodee la respuesta correcta.

- No.
- Sí.

- **En caso de que así sea ¿qué *subdimensión*/es sería/n?**

Ítem (1a, 1b, 1c...)	**Subdimensión**
________________	____
________________	____
________________	____
________________	____
________________	____
________________	____
________________	____

DIMENSIÓN 1: PROBLEMAS EN LA COMUNICACIÓN E INTERACCIÓN SOCIAL.												
***SUBDIMENSIÓN 2:* DEFICIENCIAS EN LAS CONDUCTAS COMUNICATIVAS NO VERBALES UTILIZADAS EN LA INTERACCIÓN SOCIAL**												
Opciones de respuesta: *1 (ninguna), 2 (leve), 3 (moderada) y 4 (totalmente).*												
Ítem	**Relevancia para la sub-dimensión**				**Adecuación a los objetivos**				**Respuesta diferencial en función de distintos grupos**			
2a Durante una conversación con otras personas suelo dejar largas pausas.	1	2	3	4	1	2	3	4	1	2	3	4
2b Me resulta fácil “leer entre líneas” cuando alguien me habla.	1	2	3	4	1	2	3	4	1	2	3	4
2c Soy capaz de darme cuenta si una persona que me está escuchando se aburre.	1	2	3	4	1	2	3	4	1	2	3	4
2d Levanto la voz demasiado sin darme cuenta.	1	2	3	4	1	2	3	4	1	2	3	4
2e Me resulta fácil adivinar lo que alguien está pensando o sintiendo con sólo mirarle a la cara.	1	2	3	4	1	2	3	4	1	2	3	4
2f Cuando converso con alguien mi voz tiene una entonación plana o monótona.	1	2	3	4	1	2	3	4	1	2	3	4
2g Los demás piensan que mis expresiones faciales son demasiado serias.	1	2	3	4	1	2	3	4	1	2	3	4
2h Tiendo a hablar con una voz monótona (en otras palabras, con menos cambios en la voz que la mayoría de la gente).	1	2	3	4	1	2	3	4	1	2	3	4
2i Mis expresiones faciales proporcionan información errónea a los demás acerca de cómo me siento.	1	2	3	4	1	2	3	4	1	2	3	4
2j Sé cuándo alguien no está interesado en lo que estoy diciendo.	1	2	3	4	1	2	3	4	1	2	3	4

2k	**Cuando estoy conversando me doy cuenta de cuándo es el momento de cambiar de tema.**	1	2	3	4	1	2	3	4	1	2	3	4
2l	**Me resulta difícil adivinar las intenciones de los demás.**	1	2	3	4	1	2	3	4	1	2	3	4

- **¿Considera que alguno de los anteriores ítems podría ser más adecuado para medir cualquier otra *subdimensión* distinta de esta?** Consulte la Tabla de contenido del apartado 3 de las Instrucciones para llevar a cabo el juicio de expertos y rodee la respuesta correcta.

- No.
- Sí.

- **En caso de que así sea ¿qué *subdimensión*/es sería/n?**

Ítem (2a, 2b, 2c...)	Subdimensión
________________	____
________________	____
________________	____
________________	____
________________	____
________________	____
________________	____

DIMENSIÓN 1: PROBLEMAS EN LA COMUNICACIÓN E INTERACCIÓN SOCIAL.

<u>SUBDIMENSIÓN 3:</u> DEFICIENCIAS EN EL DESARROLLO, MANTENIMIENTO Y COMPRENSIÓN DE LAS RELACIONES.

Opciones de respuesta: *1 (ninguna), 2 (leve), 3 (moderada) y 4 (totalmente).*

Ítem		Relevancia para la sub-dimensión				Adecuación a los objetivos				Respuesta diferencial en función de distintos grupos			
3a	**Cuando leo una historia, tengo facilidad para imaginarme cómo podría ser el aspecto de los personajes.**	1	2	3	4	1	2	3	4	1	2	3	4
3b	**Soy capaz de imitar las acciones o expresiones de otros cuando es socialmente apropiado hacer.**	1	2	3	4	1	2	3	4	1	2	3	4
3c	**Cuando leo una historia, encuentro difícil adivinar las intenciones de los personajes.**	1	2	3	4	1	2	3	4	1	2	3	4
3d	**Interactúo de manera apropiada con otros adultos.**	1	2	3	4	1	2	3	4	1	2	3	4
3e	**Soy una persona diplomática.**	1	2	3	4	1	2	3	4	1	2	3	4
3f	**A veces cometo el error de caminar entre dos personas que están tratando de hablar entre ella.**	1	2	3	4	1	2	3	4	1	2	3	4
3g	**Me resulta muy fácil jugar con niños/as a juegos simbólicos o inventados.**	1	2	3	4	1	2	3	4	1	2	3	4
3h	**Tengo dificultades para relacionarme con adultos que no sean de mi familia.**	1	2	3	4	1	2	3	4	1	2	3	4
3i	**Tengo dificultades para relacionarme con los miembros de mi familia.**	1	2	3	4	1	2	3	4	1	2	3	4

3j	**Me gusta estar con otras personas.**	1	2	3	4	1	2	3	4	1	2	3	4
3k	**Me cuesta imaginarme cómo sería ser otra persona.**	1	2	3	4	1	2	3	4	1	2	3	4
3l	**Disfruto estando en situaciones sociales.**	1	2	3	4	1	2	3	4	1	2	3	4

DIMENSIÓN 1: PROBLEMAS EN LA COMUNICACIÓN E INTERACCIÓN SOCIAL.
SUBDIMENSIÓN 3: ***DEFICIENCIAS EN EL DESARROLLO, MANTENIMIENTO Y COMPRENSIÓN DE LAS RELACIONES.***

Opciones de respuesta: *1 (ninguna), 2 (leve), 3 (moderada) y 4 (totalmente).*

Ítem		**Relevancia para la sub-dimensión**				**Adecuación a los objetivos**				**Respuesta diferencial en función de distintos grupos**			
3m	**Cuando era un niño/a, solía pasármelo bien jugando con otros/as niños/as a juegos simbólicos o inventados.**	1	2	3	4	1	2	3	4	1	2	3	4
3n	**Evito comenzar interacciones sociales con otros adultos.**	1	2	3	4	1	2	3	4	1	2	3	4
3o	**Me siento mucho más tenso/a en ambientes sociales que cuando estoy solo/a.**	1	2	3	4	1	2	3	4	1	2	3	4
3p	**Prefiero hacer cosas con otras personas en vez de hacerlas solo/a.**	1	2	3	4	1	2	3	4	1	2	3	4
3q	**Disfruto conociendo gente nueva.**	1	2	3	4	1	2	3	4	1	2	3	4
3r	**Cuando doy conversación es sólo para ser educado/a.**	1	2	3	4	1	2	3	4	1	2	3	4
3s	**Me río en momentos inapropiados.**	1	2	3	4	1	2	3	4	1	2	3	4

- **¿Considera que alguno de los anteriores ítems podría ser más adecuado para medir cualquier otra *subdimensión* distinta de esta?** Consulte la Tabla de contenido del apartado 3 de las Instrucciones para llevar a cabo el juicio de expertos y rodee la respuesta correcta.

- No.
- Sí.

- **En caso de que así sea ¿qué *subdimensión*/es sería/n?**

Ítem (3a, 3b, 3c...)	**Subdimensión**
______________	____
______________	____
______________	____
______________	____
______________	____
______________	____
______________	____

DIMENSIÓN 2: PATRONES RESTRICTIVOS Y REPETITIVOS DE COMPORTAMIENTO, INTERESES O ACTIVIDADES.
***<u>SUBDIMENSIÓN 1:</u>* PATRONES RESTRICTIVOS Y REPETITIVOS DE COMPORTAMIENTO, INTERESES O ACTIVIDADES.**

Opciones de respuesta: *1 (ninguna), 2 (leve), 3 (moderada) y 4 (totalmente).*

Ítem	**Relevancia para la sub-dimensión**				**Adecuación a los objetivos**				**Respuesta diferencial en función de distintos grupos**			
4a Tengo comportamientos repetitivos que los demás consideran extraños.	1	2	3	4	1	2	3	4	1	2	3	4

- **¿Considera que este ítem (4a) podría ser más adecuado para medir cualquier otra *subdimensión* distinta de esta?** Consulte la Tabla de contenido del apartado 3 de las Instrucciones para llevar a cabo el juicio de expertos y rodee la respuesta correcta.

- No.
- Sí.

- **En caso de que así sea ¿qué *subdimensión* sería?**

Subdimensión

____

____

____

____

DIMENSIÓN 2: PATRONES RESTRICTIVOS Y REPETITIVOS DE COMPORTAMIENTO, INTERESES O ACTIVIDADES.
SUBDIMENSIÓN 2: ***INSISTENCIA EN LA MONOTONÍA, EXCESIVA INFLEXIBILIDAD DE RUTINAS O PATRONES RITUALIZADOS DE COMPORTAMIENTO VERBAL O NO VERBAL.***

Opciones de respuesta: *1 (ninguna), 2 (leve), 3 (moderada) y 4 (totalmente).*

Ítem		**Relevancia para la sub-dimensión**				**Adecuación a los objetivos**				**Respuesta diferencial en función de distintos grupos**			
5a	**Me gusta planificar minuciosamente cualquier actividad en la que participo.**	1	2	3	4	1	2	3	4	1	2	3	4
5b	**Me gusta seguir una rutina minuciosa cuando trabajo.**	1	2	3	4	1	2	3	4	1	2	3	4
5c	**Prefiero hacer las cosas siempre de la misma manera.**	1	2	3	4	1	2	3	4	1	2	3	4
5d	**Las situaciones novedosas me producen ansiedad.**	1	2	3	4	1	2	3	4	1	2	3	4
5e	**Soy flexible sobre cómo deberían hacerse las cosas.**	1	2	3	4	1	2	3	4	1	2	3	4
5f	**Tiendo a ser inflexible.**	1	2	3	4	1	2	3	4	1	2	3	4
5g	**No me molesta que mi rutina diaria se vea alterada.**	1	2	3	4	1	2	3	4	1	2	3	4
5h	**Soy muy rígido/a en mis costumbres.**	1	2	3	4	1	2	3	4	1	2	3	4
5i	**Tengo más dificultad que los demás para aceptar los cambios que se producen en mi rutina.**	1	2	3	4	1	2	3	4	1	2	3	4
5j	**Tengo dificultades para enfrentarme a cambios en mi rutina.**	1	2	3	4	1	2	3	4	1	2	3	4

- **¿Considera que alguno de los anteriores ítems podría ser más adecuado para medir cualquier otra *subdimensión* distinta de esta?** Consulte la Tabla de contenido del apartado 3 de las Instrucciones para llevar a cabo el juicio de expertos y rodee la respuesta correcta.

- No.
- Sí.

- **En caso de que así sea ¿qué *subdimensión/es* sería/n?**

Ítem (5a, 5b, 5c…)	***Subdimensión***
________________	____
________________	____
________________	____
________________	____
________________	____
________________	____
________________	____

DIMENSIÓN 3: PATRONES RESTRICTIVOS Y REPETITIVOS DE COMPORTAMIENTO, INTERESES O ACTIVIDADES.
<u>SUBDIMENSIÓN 3:</u> INTERESES MUY RESTRINGIDOS Y FIJOS QUE SON ANORMALES EN CUANTO A SU INTENSIDAD O FOCO DE INTERÉS.

Opciones de respuesta: *1 (ninguna), 2 (leve), 3 (moderada) y 4 (totalmente).*

Ítem		**Relevancia para la sub-dimensión**				**Adecuación a los objetivos**				**Respuesta diferencial en función de distintos grupos**			
6a	**Me han dicho que hablo demasiado sobre ciertos temas.**	1	2	3	4	1	2	3	4	1	2	3	4
6b	**Me gusta recopilar información sobre diferentes categorías de objetos (e.g., tipos de coches, de pájaros, de trenes, de plantas, etc.)**	1	2	3	4	1	2	3	4	1	2	3	4
6c	**Pienso o hablo sobre el mismo tema una y otra vez.**	1	2	3	4	1	2	3	4	1	2	3	4
6d	**Me fascinan los números.**	1	2	3	4	1	2	3	4	1	2	3	4
6e	**A menudo la gente me dice que vuelvo una y otra vez sobre el mismo tema.**	1	2	3	4	1	2	3	4	1	2	3	4
6f	**Una vez que empiezo a pensar en algo, no puedo sacármelo de la cabeza.**	1	2	3	4	1	2	3	4	1	2	3	4

- **¿Considera que alguno de los anteriores ítems podría ser más adecuado para medir cualquier otra *subdimensión* distinta de esta?** Consulte la Tabla de contenido del apartado 3 de las Instrucciones para llevar a cabo el juicio de expertos y rodee la respuesta correcta.

- No.
- Sí.

- **En caso de que así sea ¿qué *subdimensión/es* sería/n?**

Ítem (6a, 6b, 6c…)	***Subdimensión***
________________	____
________________	____
________________	____
________________	____
________________	____
________________	____
________________	____

DIMENSIÓN 3: PATRONES RESTRICTIVOS Y REPETITIVOS DE COMPORTAMIENTO, INTERESES O ACTIVIDADES.
SUBDIMENSIÓN 4: ***HIPER O HIPORREACTIVIDAD A LOS ESTÍMULOS SENSORIALES O INTERESES INHABITUALES POR ASPECTOS SENSORIALES DEL ENTORNO.***

Opciones de respuesta: *1 (ninguna), 2 (leve), 3 (moderada) y 4 (totalmente).*

Ítem		**Relevancia para la sub-dimensión**				**Adecuación a los objetivos**				**Respuesta diferencial en función de distintos grupos**			
7a	**A menudo percibo sonidos débiles que los demás no escuchan.**	1	2	3	4	1	2	3	4	1	2	3	4
7b	**Tengo intereses sensoriales que a los demás les parecen inusuales (e.g., oler o mirar las coas de una manera especial).**	1	2	3	4	1	2	3	4	1	2	3	4
7c	**Suelo darme cuenta de detalles que otras personas pasan por alto.**	1	2	3	4	1	2	3	4	1	2	3	4
7d	**Soy excesivamente sensible a ciertos sonidos, texturas u olores.**	1	2	3	4	1	2	3	4	1	2	3	4

¿Considera que alguno de los anteriores ítems podría ser más adecuado para medir cualquier otra *subdimensión* distinta de esta? Consulte la Tabla de contenido del apartado 3 de las Instrucciones para llevar a cabo el juicio de expertos y rodee la respuesta correcta.

- No.
- Sí.

- **En caso de que así sea ¿qué *subdimensión/es* sería/n?**

Ítem (7a, 7b, 7c...)	***Subdimensión***
______________	____
______________	____

BLOQUE 2: Propiedades del test

Para finalizar nos gustaría que evaluara algunas de las propiedades del test en su conjunto (incluyendo todos los ítems y subdimensiones anteriormente mencionadas).

8.- ¿En qué grado considera que el constructo de FAA está adecuadamente representado mediante las SUBDIMENSIONES empleadas? *Rodee la opción que considere más adecuada.*

1	**2**	**3**	**4**
Muy mal representado	Mal representado	Bien representado	Muy bien representado

9.- En caso de que considere que el constructo de FAA no se encuentre adecuadamente representado por las *subdimensiones* empleadas ¿qué dimensiones o subdimensiones considera usted que deberían incluirse?

Dimensión	Subdimensiones

(En caso de que necesite espacio adicional, este cuestionario incluye un bloque de comentarios libres y una página en blanco al final para añadir las consideraciones que usted desee).

10.- ¿En qué medida cree que el constructo de FAA está adecuadamente representado mediante los ÍTEMS empleados? *Rodee la opción que considere más adecuada.*

1	**2**	**3**	**4**
Muy mal representado	Mal representado	Bien representado	Muy bien representado

11.- En caso de que considere que los ítems presentados NO representen adecuadamente el dominio de FAA, teniendo en cuenta los objetivos y la población a la que va dirigida el test, ¿qué tipo comportamientos o ítems referidos a los mismos cree usted que sería necesario añadir?

Subdimensión	Ítems/comportamientos que sería necesario añadir
Deficiencias en la reciprocidad socioemocional.	
Deficiencias en las conductas comunicativas no verbales utilizadas en la interacción social.	
Deficiencias en el desarrollo, mantenimiento y comprensión de las relaciones.	
Movimientos, utilización de objetos o habla estereotipados o repetitivos.	
Insistencia en la monotonía, excesiva inflexibilidad de rutinas o patrones ritualizados de comportamiento verbal o no verbal.	
Intereses muy restringidos y fijos que son anormales en cuanto a su intensidad o foco de interés.	

Híper o hiporreactividad a los estímulos sensoriales o interés inhabitual por aspectos sensoriales del entorno.	

(En caso de que necesite espacio adicional, este cuestionario incluye un bloque de comentarios libres y una página en blanco al final ara añadir las consideraciones que usted desee).

12.- A continuación, se le presenta cuál será el formato de respuesta para cada uno de los ítems:

Para cada pregunta, marque con una (X) el valor que mejor describe con qué frecuencia esa frase se aplica a usted según los valores de la siguiente lista:

1. Nunca	**2. Casi Nunca**	**3. Ocasionalmente**	**4. A menudo**	**5. Casi Siempre**	**6. Siempre**

¿En qué medida considera usted que el formato de respuesta es adecuado según los objetivos del test y la población a la que éste está destinado? *Rodee la opción que considere más adecuada.*

1	**2**	**3**	**4**
Muy poco adecuado	Poco adecuado	Adecuado	Muy adecuado

13.- A continuación, se le muestran las instrucciones que se incluyen en el cuestionario:

*El estudio consta de **una sola sesión** de aproximadamente **45 minutos de duración** durante la cual se cumplimentarán los cuestionarios que se acaban de señalar**, no previéndose ningún riesgo** para los participantes en el mismo. Es importante señalar que, en caso de que dé su consentimiento para participar en dicho estudio, podrá **abandonarlo en cualquier momento** sin necesidad de alegar el motivo de su abandono. En cuanto a la confidencialidad y el tratamiento de los datos personales, **sus datos no serán revelados**, y su protección y tratamiento se realizará al amparo de la Ley 15/1999, de Protección de Datos de Carácter Personal, y al Real Decreto 1729/2007 de 21 de diciembre por el que se aprueba el Reglamento de desarrollo de dicha ley.*

Lea cuidadosamente estas instrucciones antes de comenzar para marcar correctamente todas sus respuestas. Cada cuestionario consta de una escala de respuesta específica y una serie de ítems a los que deberá contestar de acuerdo a esa escala.

*No hay respuestas correctas o incorrectas. Por favor, conteste de forma sincera y exprese sus opiniones de la manera más precisa posible. Lea cada frase con atención y marque con una **(X)** la alternativa que desea. Dé una respuesta a todas las frases. Asegúrese de que marca cada respuesta en la casilla correspondiente a la misma frase y en la opción que mejor se ajuste a lo que le preguntan. NO DEJE NINGUNA PREGUNTA SIN RESPONDER.*

14. Por favor, indique en qué medida considera usted que las instrucciones del test son adecuadas según los objetivos del mismo y la población a la que está destinado.

Rodee la opción que considere más adecuada.

1	**2**	**3**	**4**
Muy poco adecuadas	Poco adecuadas	Adecuadas	Muy adecuadas

BLOQUE 3: Sugerencias, comentarios y agradecimiento

A continuación, nos gustaría facilitarle un espacio donde usted pueda expresar libremente su opinión general acerca del test, del procedimiento de juicio de expertos o de cualquier otra cosa que estime oportuna que el equipo de investigación deba tener en cuenta.

Finalmente, nos gustaría agradecer con toda sinceridad su inestimable ayuda y colaboración con el desarrollo de nuestra investigación encaminada a construir primer test en español para medir expresiones del Fenotipo Ampliado del Autismo.

El equipo investigador: Fernando Cañadas Pérez, M. Ángeles Fernández Estévez,
Marta Godoy Giménez, Antonio González Rodríguez,
Pablo Sayans Jiménez.

Datos adicionales

La información recogida en todas las preguntas que se presentan a continuación se procesará de forma independiente de las respuestas ofrecidas en el cuestionario de expertos. El objetivo del siguiente bloque es garantizar que el panel de expertos cubre todos los ámbitos profesionales relacionados con TEA y/o FAA. Además, le solicitamos información para que, de así desearlo, podamos enviarle el cuestionario en su versión final e incluirle en los agradecimientos de la publicación donde figuren los resultados del juicio de expertos.

Por favor, indique, mediante una **X**, en cuál de los siguientes ámbitos desarrolla su labor profesional y especifique, en la medida de lo posible, cuál es su perfil profesional (*solo puede escoger un ámbito*).

Ámbito sanitario.

Atención temprana (Psicólogo, Logopeda, Neuropsicólogo, Neuropediatra)

Perfil profesional: ______________________________

Psicólogo clínico (PIR)

Perfil profesional: ______________________________

Psiquiatra

Perfil profesional: ______________________________

Psicólogo Sanitario (MPGS o licenciado convalidado)

Perfil profesional: ______________________________

Ámbito social.

Actuación de los Servicios Sociales en la promoción social de la Familia

Perfil profesional: ______________________________

Asociaciones de TEA o familiares

Perfil profesional: ______________________________

Ámbito educativo

Equipos de orientación educativa

Perfil profesional: ______________________________

Pedagogía terapéutica

Perfil profesional: ______________________________

Centros específicos

Perfil profesional: ______________________________

Ámbito académico

Área y perfil investigador

En caso de que usted quiera aparecer en los agradecimientos o quiera recibir la versión final del cuestionario (este proceso se puede demorar entre uno y dos años) facilite a continuación la información de contacto donde quiera que le sea enviado.

____

____

____

Por favor, una vez cumplimentado, introduzca ese documento en el sobre adjunto y ciérrelo. Gracias.

Appendix 5. Chapter Six. Study 4.

“The Broad Autism Phenotype - International Test (BAP-IT): A Two-domain Based Test for the Assessment of the Broad Autism Phenotype

Table 1

Candidate Pool of Items for Constructing the BAP-IT, Item Codes, Spanish and English Contents

SOCIAL BAP. Persistent deficits in social communication and social interaction across multiple contexts:

SA. Deficits in social-emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions.

Item code	Spanish items	English items
SA1A1	Cuando relleno cualquier tipo de cuestionario o solicitud, prefiero hacerlo mal antes que preguntar mis dudas a la persona responsable.	When I fill out any type of questionnaire or application, I would rather do it wrong than ask the person in charge to answer any questions I may have.
SA1A2	En un restaurante, prefiero que alguno/a de mis acompañantes pida la comida por mí.	At a restaurant, I prefer one of my companions to order the food for me.
SA1A3	Me resulta difícil pedir ayuda a desconocidos/as porque **no** sé cómo iniciar la conversación.	I find it hard to ask people I do **not** know for help because I am not sure how to start the conversation.
SA1A7	En situaciones sociales, **no** se me ocurren temas de conversación.	In social situations, I **cannot**_think of any topics of conversation.
SA1B2	Cuando interactúo con otra persona, me cuesta escuchar lo que me está contando porque estoy centrado/a en lo que voy a decir después.	When I interact with another person, I find it hard to listen to what they are telling me because I am more focused on what I am going to say when they finish.
SA1B3	Cuando se me ocurre algo que decir, suelo decirlo sin importarme si la otra persona ha terminado de hablar.	When I think of something to say, I tend to say it without caring whether the other person has finished speaking.
SA1B4	Cuando tengo confianza con una persona me pongo a hablar y **no** paro, parece un monólogo mío.	When I feel comfortable with someone, I start talking and **cannot** stop; it is like my own monologue.
SA1B10[i]	Me siento cómodo/a compartiendo los turnos de palabra en una conversación.	I feel comfortable taking turns when participating in a conversation.
SA2A2	En los descansos de una actividad grupal (ej., trabajo, clase, etc.) prefiero quedarme haciendo mis cosas (ej., adelantar trabajo, leer, etc.) antes que hablar con el resto de mis compañeros/as.	During breaks in group activities (e.g., work, class, etc.), I prefer to do my own thing (e.g., get a head-start on work, read, etc.) to speaking to the other people in the group.

Item code	Spanish items	English items
SA2A3	Entre mis actividades de ocio se encuentran actividades más individuales (ej., leer, escribir, dibujar, etc.) que grupales (ej., deportes de grupo, tocar en una banda, bailar, etc.).	Out of all my leisure activities, there are more individual ones (e.g., reading, writing, drawing, etc.) than group activities (e.g., group sports, playing in a band, dance, etc.).
SA2A5	Evito participar en fiestas o eventos sociales (ej., Navidad, cumpleaños, bodas, etc.).	I avoid attending parties and social events (e.g., Christmas, birthdays, weddings, etc.).
SA2A10[i]	Me encantan los eventos sociales.	I love social events.
SA2A17	Prefiero leer o dedicar tiempo a actividades que **no** requieren compañía que salir con amigos/as.	I prefer reading or devoting time to activities that do **not** require company rather than going out with friends.
SA3A2[i]	Cuando asisto a una clase o actividad dirigida, me gusta participar.	When I attend a class or a directed activity, I like to participate.
SA3A3	Cuando entro en una sala donde hay más gente (ej., consulta del médico), espero a que alguien me salude antes de saludar yo.	When I enter a room where there are other people (e.g., doctor's office), I wait for someone else to greet me before I do so.
SA3A7	Cuando una persona desconocida me habla, me limito a responder de manera breve a sus preguntas.	When a stranger speaks to me, I only give brief answers to their questions.
SA3A8	Cuando voy en un transporte público y alguien se sienta a mi lado, finjo estar ocupado/a (ej., cojo el móvil) para que no me hablen.	When I travel on public transportation and someone sits beside me, I pretend to be busy (e.g., I take out my mobile) so they will **not** speak to me.
SA3A9	En las clases o en cualquier tipo de reunión, me sitúo detrás para que el resto de las personas reunidas **no** se fijen en mí.	In classes or any type of meeting, I sit/stand at the back so the other people gathered will not pay attention to me.
SA3A11	Las conversaciones con personas desconocidas **no** me interesan.	Conversations with people I do **not** know do **not** interest me.
SA3A13	Es difícil hablar conmigo porque no doy pie a la conversación.	It is hard to talk to me because I do **not** help the conversation develop.
SA3A14	Me pongo los cascos/auriculares (o me pondría unos si pudiera) para evitar que otras personas me hablen.	I put my headphones/earphones on (or I would if I had them) to prevent others from talking to me.

Item code	Spanish items	English items
SA3A16[i]	Saludo a mis vecinos/as y les pregunto cómo les va cuando coincido con ellos/as.	I greet my neighbours and I ask them how they are doing when I come across them.
SB. Deficits in nonverbal communicative behaviours used for social interaction, ranging, for example, from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication.		
SB1A2	Me resulta difícil saber si alguien está siendo irónico/a si **no** me lo dice directamente.	I find it difficult to tell whether someone is being ironic if they do **not** tell me directly.
SB1A7[i]	Sé cuándo otra persona me está gastando una broma basándome en su tono de voz y/o en su postura corporal.	I can tell when someone is joking with me based on their tone of voice and/or body language.
SB1B4[i]	La expresión de mi cara y mi tono de voz van acordes con lo que quiero decir.	My facial expression and my tone of voice are in agreement with what I want to say.
SB1C2	Cuando hablo, lo que digo va por un lado y lo que expreso con mis gestos por otro (**no** hay congruencia entre lo que digo y cómo lo digo).	When I speak, what I say goes in one direction and what I express with my gestures goes in another (there is **no** coherence between what I say and how I say it).
SB1C6	Mi expresión corporal parece tensa, aunque cuente algo divertido.	My body expression appears tense, even though I am talking about something fun.
SB1C7	Mi expresión corporal se mantiene tensa independientemente de lo que esté contando o a quién se lo esté contando.	My body expression remains tense regardless of what I am talking about or who I am talking to.
SB1C9	Al hablar, tengo la sensación de que mis manos y mi cuerpo no expresan lo mismo que el mensaje que trato de transmitir.	When I speak, I get the feeling that my hands and my body do not express the same message I am trying to convey.
SB2A1	Cuando converso, miro más al cuerpo que a la cara de la persona con la que hablo.	When I speak, I get the feeling that my hands and my body do **not** express the same message I am trying to convey.
SB2A2	Durante una conversación, me he dado cuenta de estar hablando con la otra persona sin mirarle los ojos.	While have a conversation, I have realised that I have been talking to the other person without looking them in the eyes.
SB2A4	Mantengo la mirada perdida (miro a la nada) durante las conversaciones.	I avoid eye contact in conversations by gazing off in another direction.

Item code	Spanish items	English items
SB2A6	Mirar directamente a los ojos de la persona con la que hablo me genera incomodidad.	Looking directly into the eyes of the person I am talking to makes me feel awkward.
SB2A7[i]	Miro a los ojos de una persona cuando me la están presentando.	When I'm introduced to someone, I look that person in the eyes.
SB2B1[i]	Devuelvo una sonrisa a quien me sonríe.	I smile back to people who smile at me.
SB2B2[i]	En una conversación casual, si la otra persona me sonríe, le devuelvo la sonrisa.	In a casual conversation, if the other person smiles, I smile back.
SB2B5	Soy indiferente a las sonrisas de cortesía (ej., cuando alguien te sonríe por educación).	I am indifferent to courtesy smiles (e.g., when someone smiles at me to be polite).
SB3A6	Evito gesticular (ej., con la cara, con las manos, etc.) en una conversación.	I avoid gestures (e.g., with my face, hands, etc.) in conversations.
SB3B1	Me cuesta entender lo que quieren decirme las personas que gesticulan mucho cuando hablan.	When people gesture a lot when they are speaking, I find it hard to understand what they want to say to me.
SB3B2	Me cuesta interpretar con claridad el significado de algunos gestos que hace la gente al hablar.	I find it hard to clearly interpret the meaning of certain gestures that people make when they speak.
SB3B3	Me resulta difícil comprender a los demás cuando gesticulan con las manos.	I find it hard to understand others when they gesture with their hands.
SB4A1	Cuando hablo, hago pocos gestos con mi cara.	When I speak, I make few gestures with my face.
SB4A9	Prefiero **no** hacer gestos con la cara cuando hablo.	I prefer **not** to make facial gestures when I am talking.
SB4B1	Cuando estoy hablando con alguien, **no** tengo claro hacia dónde mirar.	When I am talking to someone, I am **not** sure where I should look.
SB4C2	Mantengo la misma expresión facial durante toda una conversación.	I maintain the same facial expression throughout an entire conversation.
SB4C3	Me han dicho que mi rostro es inexpresivo.	I have been told I have an expressionless face.
SB4C6	Mi cara es poco expresiva tanto cuando cuento un chiste como cuando doy una mala noticia.	My face is not very expressive, both when I tell a joke and when I have to tell some bad news.

Item code	Spanish items	English items
SB4C7	Por la expresión de mi cara, parece que estuviera enfadado/a.	My usual facial expression makes people think I am angry.
SC. Deficits in developing, maintaining, and understanding relationships, ranging, for example, from difficulties adjusting behaviour to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to absence of interest in peers.		
SC1A2	Encuentro difícil amoldar mi comportamiento a lo que se exige en algunos contextos (ej., un entierro o un acto religioso, etc.).	I find it hard to adjust my behaviour to what is required in certain contexts (e.g., a funeral or religious ceremony, etc.).
SC1A3	Evito ir a ciertos lugares (ej., hospitales, iglesias, etc.) porque **no** sé muy bien cómo comportarme.	I avoid going to certain places (e.g., hospitals, churches, etc.) because I am **not** sure how to behave.
SC1B1	Hablo de forma monótona.	I have a monotonous way of speaking.
SC1B3	Me cuesta controlar mi tono de voz si **no** estoy atento/a a ello.	It is hard for me to control my tone of voice if I am **not** paying attention to it.
SC1B4	Me dicen que mi tono de voz es plano.	I have been told I have a flat tone of voice
SC1B5	Mis amigos/as se quejan de que hablo demasiado bajo en los eventos sociales.	My friends complain that I speak too quietly at social events
SC2B1	Cuando alguien me cuenta un problema que le ha afectado emocionalmente, me cuesta conectar con el sentimiento de esa persona.	When someone tells me a problem that has affected them emotionally, it is hard for me to relate to the feelings of that person.
SC3A2[i]	Hago amigos/as en cualquier sitio.	I can make friends wherever I go.
SC3A4	Me resulta difícil hacer nuevos amigos/as.	It is hard for me to make new friends.
SC3B1	Estoy más cómodo/a conversando con mis amigos/as de toda la vida que conociendo a gente nueva.	I feel more comfortable being with the friends I have always known than meeting new people.
SC3B2	Me divierto más estando a solas con mi mejor amigo/a que cuando hay más gente.	I have more fun when I am alone with my best friend than when I am with a group of people.
SC3B5	Me molesta que mis amigos/as inviten a personas que **no** conozco a nuestras quedadas.	I get annoyed when my friends invite people I do **not** know to our gatherings.
SC3B6	Prefiero **no** relacionarme con los/as conocidos/as de mis amigos/as.	I prefer **not** to interact with the people my friends know.

Item code	Spanish items	English items
SC3B7	Prefiero quedar con mis amigos/as de toda la vida que conocer gente nueva.	I prefer to get together with the friends I have always know than to meet new people.
SC4A4	Me siento más cómodo/a trabajando solo/a que en grupo.	I am more comfortable working alone than in a group.
SC4A5	Participar en actividades grupales me supone un sacrificio.	Participating in group activities takes a lot of effort on my part.
SC4B3	Disfruto de mi soledad.	I enjoy my solitude.
SC4B4	Disfruto más de mi soledad que de la compañía de otras personas.	I enjoy my solitude more than the company of other people.
SC4B10	Prefiero hacer actividades solo/a que en grupo.	I prefer to do activities alone than in a group.
SC4B12	Soy una persona solitaria.	I am a loner.
SC4C1[i]	Disfruto en situaciones sociales donde las demás personas están disfrutando.	I enjoy social situations in which the other people are having a good time.
SC4D1	Cuando acudo a un evento social donde **no** conozco a todo el mundo, me pone nervioso/a **no** saber qué decir ni qué hacer.	When I attend a social event where I do **not** know everyone, I get worried about **not** knowing what to say or do.
SC4D5	Me preocupa lo que la gente piensa de mí cuando voy a reuniones donde **no** conozco a todo el mundo.	I worry about what people think about me when I go to meetings where I do **not** know everyone
SC4D7	Tiendo a pensar que **no** le voy a caer bien a la gente antes de conocerla.	I tend to think that people are **not** going to like me before I meet them.
RIRE BAP. Restricted, repetitive patterns of Behaviour, interests, or activities, as manifested by at least two of the following, currently or by history: *RA. Stereotyped or repetitive motor movements, use of objects, or speech (e.g., simple motor stereotypies, lining up toys or flipping objects, echolalia, idiosyncratic phrases).*		
RA1A2	Cuando estoy sentado/a, me gusta balancear mi cuerpo adelante y atrás.	When I am sitting down, I like to rock my body back and forth.
RA1A3	Cuando me pongo nervioso/a, NO puedo controlar ciertos movimientos repetitivos (ej., tocarme la cabeza, etc.).	When I get nervous (anxious), I cannot help making certain repetitive movements (e.g., touching my head, etc.).
RA1A5	Durante una conversación, muevo las manos sin sentido comunicativo.	During a conversation, I move my hands randomly, without any communicative meaning.
RA1A6	Hago gestos repetitivos sin propósito.	I make repetitive gestures without any meaning.

Item code	Spanish items	English items
RA1A7	Repito algunos movimientos como tocarme la cabeza, girar la mano, mover una o ambas piernas, etc.	I make some repetitive movements, such as touching my head, turning my hand, moving one or both legs, etc.
RA1A8	Siento la necesidad de mover mis manos cuando no las estoy usando.	I feel the need to move my hands when I am not using them.
RA2A2	De pequeño/a pasaba más tiempo ordenando mis peluches, juguetes, maquetas, etc., que jugando con ellos/as.	When I was little, I spent more time organising my dolls, toys, models, etc. than playing with them.
RA2A3	Disfruto clasificando cosas.	I enjoy classifying things.
RA2A5	Me encanta clasificar cosas siguiendo un criterio (ej., primero por colores después, por formas y, finalmente, por usos).	I love to classify things according to some criteria (e.g., first by colours, then by shapes and, finally, by uses).
RA2A8	Me gusta pasar mi tiempo libre ordenando mis cosas.	I like to spend my free time organising my things.
RA2A10	Mientras hablo, alineo, ordeno o clasifico las cosas que tengo al alcance de mi mano de forma casi involuntaria.	While I am speaking, I align, organise or classify objects within reach almost involuntarily.
RA3A1	Tengo muletillas genéricas que repito en cualquier situación.	There are certain generic phrases that I repeat in any situation.
RA3A2	Al conversar, hago uso de muletillas (palabras y sonidos recurrentes que repito por hábito).	When having a conversation, I use certain catchphrases (words and sounds that I repeat out of habit).
RA3A3	Cuando estoy nervioso/a y a solas, digo en voz alta lo primero que me viene a la cabeza.	When I am nervous (anxious) and alone, I say out loud the first thing that comes to mind.
RA3A6	Cuando escucho una frase que me gusta, la repito literalmente en otras ocasiones, aunque **no** se ajuste al contexto	When I hear a phrase that I like, I repeat it verbatim on other occasions even though it does **not** fit the context.
RA3A8	Repito frases de películas, series o canciones, aunque **no** sea para expresar algo concreto.	I repeat lines from films, series or songs, even if it is **not** to express anything in particular.
RA4A5	Me han dicho que hablo demasiado sobre mí mismo/a.	I have been told I speak too much about myself.

Item code	Spanish items	English items
RA4B2	Le doy un sentido propio (mi propio significado) a ciertas expresiones que utilizo al hablar.	I give my own meaning to certain expressions that I use when speaking.
RA4B3	El significado que le doy a algunas frases y/o expresiones **no** coincide con el que le dan otras personas.	The meaning I give to some phrases and/or expressions does **not** match with the meaning used by other people.
RA4B5	Uso una parte o característica de un objeto para referirme al objeto completo (ej., "alas" para referirme a un avión, "suave" para referirme a ciertos animales suaves, etc.).	I use a part or a characteristic of an object to refer to the entire object (e.g., "wings" to refer to a plane, "smooth" to refer to certain smooth animals, etc.).
RA4B6	Copio expresiones de otras personas y les doy un significado propio.	I copy the expressions used by other people and I give them my own meaning.
RB. Insistence on sameness, inflexible adherence to routines, or ritualized patterns or verbal nonverbal behaviour (e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take same route or eat food every day).		
RB1A5	Me pone nervioso/a que me cambien una cita a última hora.	I get anxious when an appointment is changed at the last minute.
RB1A8	Si he planeado ir a un sitio y, cuando llego, está cerrado, prefiero volverme a mi casa antes que buscar un plan alternativo.	If I have planned to go somewhere and, when I arrive, it is closed, I would rather go back home than think of an alternative plan.
RB1A9	Si tengo que seguir una receta, plan de trabajo o protocolo y no puedo realizar uno de los pasos, prefiero dejar de hacerlo.	If I have to follow a recipe, work plan or protocol and I am unable to carry out one of the steps, I prefer to stop entirely.
RB1A11[i]	Soy flexible ante cualquier cambio de planes inesperado.	I am flexible when it comes to any unexpected changes in plans.
RB1A12[i]	Me adapto bien a los cambios de planes inesperados.	I adapt well to unexpected changes in plans.
RB2A1	Aunque lo sepa con suficiente antelación, cambiar un trayecto cotidiano me genera nerviosismo.	A change in one of my daily routes makes me feel anxious, even if I am aware of it well in advance.
RB2A2	Me cuesta trabajo aceptar que el local (ej., gimnasio, librería, centro de ocio, etc.) al que solía ir a pasar mi tiempo libre haya cerrado.	It is hard for me to accept that an establishment where I used to spend my free time has closed (e.g., gym, book shop, recreation centre, etc.).

Item code	Spanish items	English items
RB2A6	Me incomoda cuando los negocios a los que voy habitualmente cierran por vacaciones.	I get annoyed when the businesses I habitually go to close for holidays.
RB2B4	Mi miedo al cambio hace que rechace opciones que podrían mejorar mi vida (ej., cambios de trabajo, estudio, domicilio, etc.).	My fear of change leads me to reject options that could possibly improve my life (e.g., changing jobs, studies, residence, etc.).
RB2B6	Los cambios grandes en mi vida (ej., de casa, trabajo, centro de estudios, etc.) me generan preocupación.	big changes in my life (e.g., home, job, school, etc.) make me worry.
RB2B7	Los cambios importantes en mi vida (ej., cambiar de trabajo, escuela o ciudad, etc.) me generan preocupación que dura más de un día.	Important changes in my life (e.g., changing jobs, schools or cities, etc.) make me worry, and the feeling lasts for more than a day.
RB3A3	Cuando alguien me lleva la contraria, siento que mi deber es hacerle ver a esa persona que se está equivocando.	When someone disagrees with me, I feel it is my duty to make them see that they are wrong.
RB3A5	En una conversación, mantengo mis ideas, aunque sepa que son incorrectas.	In a conversation, I stick to my ideas, even if I know they are wrong.
RB3A8	Me es difícil dejarme convencer por las opiniones de los/as demás.	It is hard for me to be convinced by the opinions of others.
RB3B2	Antes de salir de casa, me preparo de forma concreta para evitar salir con la sensación de haber olvidado algo.	Before leaving home, I have a specific way of preparing to avoid leaving with the feeling I have forgotten something.
RB3B3	Aunque haya herramientas o procesos mejores, considero que soy más eficiente si hago las cosas como yo estoy acostumbrado/a.	Although there may be better tools or processes, I feel I am more efficient if I do things the way I am used to.
RB3B12	Para desplazarme de un punto a otro, me gusta ir por los sitios por los que estoy habituado/a, aunque haya mejores rutas.	When I go from one place to another, I like to go the way I am accustomed to, even though there may be better routes.
RB3B13	Prefiero ir a las tiendas de toda la vida que probar una tienda nueva con mejores ofertas o productos.	I prefer to go to the same shops as always rather than trying a new shop with better offers or products.

Item code	Spanish items	English items
RB4A1	Cuando escucho una frase graciosa en una película, la utilizo para saludar a todo el mundo (ej., familiares, amigos/as cercanos/as, conocidos/as, desconocido/as, etc.).	When I hear a funny phrase in a film, I use it to greet everyone (e.g., relatives, close friends, acquaintances, strangers, etc.).
RB4A2	Independientemente de la situación, saludo de la manera en que me siento más cómodo/a.	Regardless of the context, I use my own ways of greeting (e.g., sayings, set phrases) instead of conventional greetings.
RB4A6	Me incomodan las situaciones sociales en las que hay que saludar de un modo determinado y no como yo suelo hacerlo.	I feel uncomfortable in social situations in which I have to greet people in a specific way and not the way I usually do.
RB4A13	Tengo un saludo típico o propio que uso en cualquier tipo de situación.	I use a typical greeting or my own greeting in any type of situation.
RB4A15	Independientemente del contexto, uso mis propias formas de saludar (ej., frases hechas, muletillas) antes que saludos convencionales.	Regardless of the context, I use my own ways of greeting (e.g., sayings, set phrases) instead of conventional greetings.
RB5A2	Como los mismos alimentos.	I am in the habit of eating the same foods.
RB5A3	Cuando me invitan a alguna celebración o evento, me siento mal si **no** puedo comer lo que normalmente como.	When I am invited to a celebration or event, I feel uncomfortable if I **cannot** eat what I normally do.
RB5A4	Evito probar comidas nuevas.	I avoid trying new foods.
RB5A9	Me molesta si **no** puedo comer lo que tenía previsto.	I get annoyed if I **cannot** eat what I had planned.
RB5B8	Si sé que van a servir comida exótica en una celebración (ej., cumpleaños, boda, etc.) prefiero **no** ir.	If I know that exotic food is going to be served at a celebration (e.g., a birthday, a wedding, etc), I prefer **not** to go.
RC. Highly restricted, fixated interests that are abnormal in intensity or focus (e.g., strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interest).		
RC1A2	Afronto mal las críticas.	I take criticism badly.
RC1A4	Cuando me critican por haber hecho algo de forma incorrecta, me siento fatal.	When I am criticized for doing something incorrectly, I feel terrible.
RC1A5	Cuando me critican, **no** puedo evitar enfadarme.	When I am criticized, I **cannot** help getting angry.

Item code	Spanish items	English items
RC1A7	Me molesta cuando me hacen un comentario negativo sobre mis gustos o aficiones.	I get annoyed when someone makes a negative comment about my likes or hobbies.
RC1A9	Percibo los comentarios sobre mi trabajo como un ataque a mi persona.	I take comments about my work as a personal attack against me as a person.
RC1A11	Soy sensible a las críticas.	I am sensitive to criticism.
RC2A3	Me encantan los objetos y juegos mecánicos donde el movimiento de unas partes afecta al de otras (ej., trenes, coches, objetos en equilibrio, bolas que chocan, etc.).	I love objects and mechanical games in which the movement of certain parts affects others (e.g., trains, cars, balance toys, pendulum balls, etc.).
RC2A4	Me gusta pasar el tiempo realizando actividades mecánicas como hacer puzles o rompecabezas, hacer estructuras de bloques, jugar al solitario, etc.	I enjoy spending time on mechanical activities, such as doing puzzles, making structures with blocks, playing solitaire, etc.
RC2A6	Pierdo la noción del tiempo cuando lo dedico a aficiones manuales (ej., manualidades, mecánica, carpintería, bricolaje, etc.).	I lose track of time when spend it on manual activities (e.g., arts and crafts, mechanical activities, carpentry, DIY, etc.).
RC2B5	Me fijo en todos los números con los que me encuentro (ej., las matrículas de los coches, las fechas, los números de carnés, etc.).	I pay close attention to all the numbers I come across (e.g., number plates on cars, dates, identity card numbers, etc.).
RC2C2	Guardo con cariño objetos cotidianos que la gente no suele valorar.	I save everyday objects with great care, which other people usually place little value on.
RC2C4	Me preocupo cuando se me rompe un objeto sin valor aparente (solo tiene valor para mí y no para los demás).	I get worried when an object which would appear to have no apparent value is broken (it is only of value to me and not to others).
RC2C5	Me preocupo cuando se me pierde un objeto sin valor aparente (solo tiene valor para mí y no para los demás).	I get worried when I cannot find an object which would appear to have no apparent value (it is only of value to me and not to others).
RC2C7	Cuando se me pierde o rompe algo me preocupo más que el resto de la gente.	When I lose or break something, I worry more than other people.

Item code	Spanish items	English items
RD. Hyper- or hyporeactivity to sensory input or unusual interests in sensory aspects of the environment (e.g., apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement).		
RD1A2	Me molesta menos que al resto de la gente coger cosas muy calientes con las manos.	Grabbing hot things with my hands bothers me less than it does other people.
RD1A3	Me resulta difícil saber cuándo algo está muy caliente.	I find it hard to tell when something is very hot.
RD1A5	Me tomo la comida sin importarme si está muy caliente.	I eat food without caring whether it is very hot.
RD1A10	Tengo problemas para escoger la ropa según el clima porque **no** me afectan los cambios en la temperatura.	I have a hard time choosing clothes according to the weather because changes in temperature do **not** affect me.
RD2A1	El ruido me bloquea, es decir, **no** me deja pensar con propiedad y necesito alejarme de él.	Noise completely distracts me, meaning it does **not** let me think properly and I need to get away from it.
RD2A2	Existen sonidos específicos que me molestan.	There are specific sounds that bother me.
RD2A5	Evito los sonidos fuertes porque me ponen nervioso/a.	I avoid loud sounds because they make me anxious.
RD2A6	Me molestan los sonidos estridentes (ej., agudos o desapacibles) más que al resto de las personas.	High-pitched sounds (e.g., shrill or unpleasant) bother me more than they do other people.
RD2A8	Evito ir a sitios donde sé que va a haber cierto tipo de sonidos que no me gustan.	I avoid going to places where I know there is going to be a certain type of sound that I do not like.
RD2A10	Soy más sensible que los/as demás a los sonidos fuertes.	I am more sensitive to loud sounds than others.
RD3A1	Cuando algo me resulta agradable al tacto, siento que una fuerza superior a mí me impulsa a tocarlo.	When I find something pleasant to the touch, I feel a power beyond me which drives me to touch it.
RD3A3	Cuando algo tiene un tacto agradable, NO puedo parar de tocarlo.	When something has a pleasant touch, I cannot stop touching it.
RD3A4	Cuando camino por la calle voy tocando todas las cosas que están a mi alcance.	I feel the need to touch all the fabrics that are <u>around me</u>.
RD3A7	Me detengo a oler cada uno de los alimentos que me voy a comer.	When I am walking in the street, I touch everything that is within my reach.

Item code	Spanish items	English items
RD3A10	Me gusta oler los objetos nuevos.	I like to smell new objects.
RD3A12	Siento la necesidad de tocar todos los tejidos que encuentro a mi alcance.	I feel the need to touch all the fabrics that are around me.
RD3A14	Tocar determinados objetos de forma repetida me relaja.	Touching specific objects repeatedly relaxes me.
RD4A2	Me encantan los juegos de luces y sombras.	I love light and shade art.
RD4A3	Me fascina el movimiento oscilatorio de ciertos objetos (ej., un péndulo).	I am fascinated by the oscillating movement of certain objects (e.g., a pendulum).
RD4A4	Me gusta pasar el rato mirando por la ventana y viendo la gente y los coches pasar.	I enjoy passing time looking out the window and watching people and cars go by.
RD4A6	Me quedo embelesado/a (fascinado/a) mirando las luces que cambian de colores.	I become captivated (fascinated) by lights that change colours.
RD4A7	Me quedo absorto/a (ensimismado/a) mirando luces en las que los demás **no** parecen reparar.	I become absorbed (engrossed) by lights I see which other people do **not** even seem to notice.

Note. Table contains all the items from the candidate pool organized in BAP subdomains.
Negative expressions in sentences were presented in bold lettering to participants.
[i] Items with reversed content.

Table 2

Alternative Pool of Items for Constructing the BAP-IT, Item Codes, Spanish and English Contents

Item code	Spanish items	English items
SOCIAL BAP. Persistent deficits in social communication and social interaction across multiple contexts:		
SA. Deficits in social-emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions.		
BAPQ36[i]	Disfruto charlando con la gente.	I enjoy chatting with people.
SB. Deficits in nonverbal communicative behaviours used for social interaction, ranging, for example, from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication.		
SRS12[i]	Soy capaz de expresar mis sentimientos a otras personas.	I am able to communicate my feeling to others.
SRS60	Los demás piensan que soy emocionalmente distante y que no muestro mis sentimientos.	Other people think I am emotionally distant and do not show my feelings.
SC. Deficits in developing, maintaining, and understanding relationships, ranging, for example, from difficulties adjusting behaviour to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to absence of interest in peers.		
AQ44	Disfruto estando en situaciones sociales.	I enjoy social occasions.
AQ47	Disfruto conociendo gente nueva.	I enjoy meeting new people.
BAPQ1[i]	Me gusta estar con otras personas.	I like being around other people.
RSR64	Me siento mucho más tenso/a en ambientes sociales que cuando estoy solo/a	I am much tenser in social settings than when I am by myself.
RIRE BAP. Restricted, repetitive patterns of Behaviour, interests, or activities, as manifested by at least two of the following, currently or by history:		
RA. Stereotyped or repetitive motor movements, use of objects, or speech (e.g., simple motor stereotypies, lining up toys or flipping objects, echolalia, idiosyncratic phrases).		
———	———	———
RB. Insistence on sameness, inflexible adherence to routines, or ritualized patterns or verbal nonverbal behaviour (e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take same route or eat food every day).		
———	———	———
RC. Highly restricted, fixated interests that are abnormal in intensity or focus (e.g., strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interest).		

Item code	Spanish items	English items
AQ39	A menudo la gente me dice que vuelvo una y otra vez sobre el mismo tema.	People often tell me that I keep going on and on about the same thing.
BAPQ17	Me han dicho que hablo demasiado sobre ciertos temas.	I have been told that I talk too much about certain topics.
SRS31	Una vez que empiezo a pensar en algo, **no** puedo sacármelo de la cabeza.	I **can't** get my mind off something once I start thinking about it.
RD. Hyper- or hyporeactivity to sensory input or unusual interests in sensory aspects of the environment (e.g., apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement).		
—	—	—

Note. Table contains items from the alternative pool (Godoy-Giménez et al., 2021a) organized in BAP subdomains.
Negative expressions in sentences were presented in bold lettering to participants.
[i] Items with reversed content.

Table 3

Number of Items in Each Subdomain from Each Pool

SOCIAL-BAP: Persistent deficits in social communication and social interaction across multiple contexts.	BAP-IT candidate items (n 72)	Alternative pool of items (n 7)
A. Deficits in social-emotional reciprocity	22	1
B. Deficits in non-verbal communicative behaviours used for social interaction	26	2
C. Deficits in developing, maintaining, and understanding relationships	24	4
RIRE-BAP: Restricted and/or repetitive patterns of behaviours, interests, or activities.	**BAP-IT candidate items (n 86)**	**Alternative pool of items (n 3)**
A. Stereotyped or repetitive motor movements, use of objects, or speech	21	______
B. Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behaviour	28	______
C. Highly restricted, fixated interests	15	3
D. Hyper- or hyporeactivity to sensory input or unusual interest in sensory aspects of the environment	22	______

Table 4

Descriptive Results of all the Items Included in the Results of the Present Manuscript

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
BAP-IT Candidate Items						
SA1A1						
ES1	1.92 (.04)	2	1.16	1, 6	1.48	1.89
ES2	1.71 (.05)	1	1.02	1, 6	1.93	4.26
EN	2.08 (.06)	2	1.24	1, 6	1.22	0.87
SA1A2						
ES1	2.11 (.04)	2	1.33	1, 6	1.27	0.96
ES2	2.08 (.06)	2	1.20	1, 6	1.24	1.13
EN	2.06 (.07)	1	1.40	1, 6	1.29	0.78
SA1A3						
ES1	2.72 (.05)	2	1.44	1, 6	0.70	-0.36
ES2	2.63 (.06)	2	1.30	1, 6	0.81	0.19
EN	2.99 (.07)	3	1.43	1, 6	0.44	-0.51
SA1A7						
ES1	2.82 (.04)	3	1.14	1, 6	0.71	-0.22
ES2	2.58 (.06)	2	1.20	1, 6	0.92	0.59
EN	2.83 (.06)	3	1.33	1, 6	0.47	-0.28
SA1B2						
ES1	2.51 (.04)	2	1.17	1, 6	0.71	0.16
ES2	2.38 (.05)	2	1.13	1, 6	0.94	0.87
EN	2.82 (.06)	3	1.32	1, 6	0.52	-0.36
SA1B3						
ES1	2.42 (.04)	2	1.11	1, 6	0.95	0.77
ES2	2.18 (.05)	2	1.00	1, 6	0.91	0.84
EN	2.59 (.06)	2	1.27	1, 6	0.76	0.20
SA1B4						
ES1	2.73 (.04)	3	1.31	1, 6	0.64	-0.22
ES2	2.74 (.06)	3	1.26	1, 6	0.73	0.11
EN	2.91 (.06)	3	1.33	1, 6	0.54	-0.24
SA1B10[i]						
ES1	4.35 (.04)	5	1.24	1, 6	-0.53	-0.34
ES2	4.53 (.06)	5	1.26	1, 6	-0.65	-0.38
EN	3.94 (.06)	4	1.40	1, 6	-0.32	-0.65
SA2A2						
ES1	2.63 (.05)	2	1.39	1, 6	0.79	-0.13
ES2	2.48 (.06)	2	1.26	1, 6	1.01	0.63
EN	3.46 (.06)	3	1.37	1, 6	0.28	-0.60

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
SA2A3						
ES1	3.74 (.05)	4	1.50	1, 6	-0.05	1.05
ES2	3.41 (.07)	3	1.41	1, 6	0.27	-0.79
EN	3.81 (.06)	4	1.38	1, 6	-0.05	-0.75
SA2A5						
ES1	2.38 (.05)	2	1.39	1, 6	0.90	-0.04
ES2	2.01 (.06)	2	1.19	1, 6	1.33	1.38
EN	2.92 (.07)	3	1.53	1, 6	0.48	-0.70
SA2A10[i]						
ES1	3.68 (.05)	4	1.45	1, 6	0.06	-0.90
ES2	3.86 (.06)	4	1.35	1, 6	-0.08	-0.85
EN	3.42 (.06)	3	1.36	1, 6	0.12	-0.63
SA2A17						
ES1	3.03 (.04)	3	1.29	1, 6	0.50	-0.27
ES2	2.75 (.06)	3	1.18	1, 6	0.67	0.36
EN	3.42 (.06)	3	1.74	1, 6	0.34	-0.40
SA3A2[i]						
ES1	3.56(.05)	3	1.44	1, 6	-0.10	-0.95
ES2	3.85 (.07)	4	1.41	1, 6	-0.05	-1.05
EN	2.55 (.06)	3.5	1.33	1, 6	0.02	-0.55
SA3A3						
ES1	2.61 (.05)	2	1.45	1, 6	0.85	-0.21
ES2	2.09 (.06)	2	1.26	1, 6	1.28	1.12
EN	2.82 (.06)	3	1.39	1, 6	0.62	-0.27
SA3A7						
ES1	3.34 (.05)	3	1.40	1, 6	0.38	-0.76
ES2	2.94 (.06)	3	1.30	1, 6	0.56	-0.35
EN	3.25 (.06)	3	1.25	1, 6	0.36	-0.26
SA3A8						
ES1	3.13 (.05)	3	1.51	1, 6	0.47	-0.79
ES2	2.69 (.06)	2	1.34	1, 6	0.76	-0.06
EN	3.14 (.07)	3	1.51	1, 6	0.41	-0.76
SA3A9						
ES1	3.11 (.05)	3	1.45	1, 6	0.38	-0.77
ES2	2.74 (.07)	3	1.39	1, 6	-0.43	3.31
EN	3.31 (.07)	3	1.52	1, 6	0.28	-0.84
SA3A11						
ES1	2.89 (.04)	3	1.34	1, 6	0.60	-0.27

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
ES2	2.57 (.06)	2	1.21	1, 6	0.85	0.43
EN	2.77 (.06)	3	1.29	1, 6	0.63	-0.00
SA3A13						
ES1	2.26 (.04)	2	1.31	1, 6	1.02	0.40
ES2	1.93 (.05)	2	1.02	1, 6	1.30	1.96
EN	2.49 (.06)	2	1.34	1, 6	0.84	0.11
SA3A14						
ES1	2.46 (.05)	2	1.41	1, 6	0.91	0.07
ES2	2.15 (.06)	2	1.27	1, 6	1.20	1.08
EN	2.56 (.07)	2	1.55	1, 6	0.69	-0.54
SA3A16[i]						
ES1	3.59 (.05)	3	1.48	1, 6	0.02	-0.92
ES2	4.01 (.07)	4	1.44	1, 6	-0.18	-1.06
EN	4.01 (.07)	4	1.52	1, 6	-0.25	-0.92
SB1A2						
ES1	2.18 (.03)	2	1.06	1, 6	1.24	1.93
ES2	2.23 (.05)	2	1.09	1, 6	1.03	0.96
EN	2.54 (.06)	2	1.23	1, 6	0.82	0.29
SB1A7[i]						
ES1	4.33 (.04)	5	1.23	1, 6	0.56	-0.39
ES2	4.40 (.05)	5	1.12	1, 6	-0.45	-0.55
EN	4.07 (.05)	4	1.15	1, 6	-0.30	-0.38
SB1C2						
ES1	1.71 (.03)	2	0.89	1, 6	1.74	4.24
ES2	1.64 (.04)	1	0.87	1, 6	1.74	3.90
EN	2.12 (.06)	2	1.21	1, 6	1.12	0.85
SB1B4[i]						
ES1	4.59 (.04)	5	1.16	1, 6	0.69	-0.04
ES2	4.72 (.05)	5	1.15	1, 6	-0.98	0.82
EN	4.10 (.05)	4	1.16	1, 6	-0.31	-0.34
SB1C6						
ES1	2.16 (.04)	2	1.14	1, 6	1.07	0.91
ES2	1.96 (.05)	2	1.07	1, 6	1.23	1.41
EN	2.56 (.06)	2	1.34	1, 6	0.68	-0.22
SB1C7						
ES1	2.10 (.03)	2	1.07	1, 6	1.13	1.26
ES2	1.89 (.05)	2	1.03	1, 6	1.34	1.91
EN	2.47 (.06)	2	1.35	1, 6	0.74	-0.21

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
SB1C9						
ES1	1.96 (.03)	2	1.07	1, 6	1.36	1.87
ES2	1.78 (.05)	2	0.97	1, 6	1.64	3.17
EN	2.16 (.06)	2	1.20	1, 6	1.05	0.73
SB2A1						
ES1	2.48 (.04)	2	1.30	1, 6	0.95	0.27
ES2	2.07 (.06)	2	1.18	1, 6	1.24	1.30
EN	2.61 (.06)	2	1.30	1, 6	0.65	-0.15
SB2A2						
ES1	2.70 (.04)	2	1.32	1, 6	0.68	-0.24
ES2	2.34 (.06)	2	1.18	1, 6	0.91	0.58
EN	2.70 (.06)	3	1.32	1, 6	0.64	-0.08
SB2A4						
ES1	2.21 (.04)	2	1.23	1, 6	0.98	0.36
ES2	1.81 (.05)	2	0.99	1, 6	1.21	2.68
EN	2.68 (.07)	2	1.48	1, 6	0.63	-0.54
SB2A6						
ES1	2.74 (.05)	2	1.46	1, 6	0.77	-0.28
ES2	2.27 (.06)	2	1.23	1, 6	0.96	0.43
EN	2.94 (.07)	3	1.50	1, 6	0.49	-0.56
SB2A7[i]						
ES1	4.64 (.04)	5	1.37	1, 6	0.88	-0.07
ES2	4.97 (.06)	5	1.26	1, 6	-1.08	0.07
EN	4.08 (.06)	4	1.38	1, 6	-0.36	-0.62
SB2B1[i]						
ES1	4.87 (.04)	5	1.17	1, 6	1.00	0.39
ES2	5.21 (.05)	6	1.02	2, 6	-1.28	0.89
EN	4.83 (.05)	5	1.14	1, 6	-0.70	-0.21
SB2B2[i]						
ES1	4.86 (.04)	5	1.20	1, 6	-0.92	0.42
ES2	5.13 (.05)	5	1.06	1, 6	-1.13	0.39
EN	4.61 (.06)	5	1.22	1, 6	-0.62	-0.23
SB2B5						
ES1	2.07 (.04)	2	1.14	1, 6	1.33	1.67
ES2	1.75 (.05)	1	0.99	1, 6	1.55	2.55
EN	2.44 (.06)	2	1.34	1, 6	0.81	0.01
SB3A6						
ES1	2.00 (.03)	2	1.07	1, 6	1.43	2.08

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
ES2	2.11 (.05)	2	1.15	1, 6	1.36	1.69
EN	2.55 (.06)	2	1.26	1, 6	0.80	0.29
SB3B1						
ES1	2.13 (.03)	2	0.98	1, 6	1.17	1.95
ES2	2.05 (.05)	2	0.97	1, 6	1.18	1.88
EN	2.42 (.05)	2	1.15	1, 6	0.82	0.45
SB3B2						
ES1	2.28 (.03)	2	1.03	1, 6	1.03	1.23
ES2	2.15 (.04)	2	0.93	1, 6	0.97	1.20
EN	2.54 (.05)	2	1.10	1, 6	0.71	0.61
SB3B3						
ES1	1.96 (.03)	2	0.95	1, 6	1.42	2.96
ES2	1.88 (.04)	2	0.84	1, 6	1.12	2.16
EN	2.40 (.05)	2	1.07	1, 6	0.81	0.78
SB4A1						
ES1	2.41 (.04)	2	1.20	1, 6	0.95	0.41
ES2	2.27 (.06)	2	1.20	1, 6	1.16	1.05
EN	2.86 (.06)	3	1.26	1, 6	0.48	-0.24
SB4A9						
ES1	2.19 (.04)	2	1.24	1, 6	1.29	1.24
ES2	1.99 (.05)	2	1.14	1, 6	1.40	1.93
EN	2.45 (.06)	2	1.21	1, 6	0.69	0.06
SB4B1						
ES1	2.44 (.05)	2	1.43	1, 6	0.95	0.04
ES2	2.10 (.06)	2	1.18	1, 6	1.16	1.00
EN	2.71 (.07)	3	1.40	1, 6	0.59	-0.36
SB4C2						
ES1	1.93 (.04)	2	1.08	1, 6	1.41	1.75
ES2	2.02 (.05)	2	1.16	1, 6	1.36	1.61
EN	2.49 (.05)	2	1.14	1, 6	0.67	0.07
SB4C3						
ES1	1.60 (.03)	1	1.00	1, 6	2.07	4.60
ES2	1.46 (.04)	1	0.87	1, 6	2.38	6.39
EN	1.86 (.06)	1	1.25	1, 6	1.38	0.99
SB4C6						
ES1	1.91 (.03)	2	1.07	1, 6	1.40	1.91
ES2	1.64 (.05)	1	1.00	1, 6	1.96	4.18
EN	2.20 (.06)	2	1.25	1, 6	0.94	0.13

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
SB4C7						
ES1	2.51 (.04)	2	1.27	1, 6	0.79	0.04
ES2	2.36 (.06)	2	1.21	1, 6	0.85	0.48
EN	2.27 (.06)	2	1.30	1, 6	0.79	-0.19
SC1A2						
ES1	2.06 (.04)	2	1.23	1, 6	1.32	1.37
ES2	1.80 (.05)	1	1.13	1, 6	1.88	3.76
EN	2.08 (.06)	2	1.31	1, 6	1.16	0.53
SC1A3						
ES1	2.07 (.04)	2	1.23	1, 6	1.25	1.23
ES2	1.69 (.05)	1	1.00	1, 6	1.90	4.08
EN	2.25 (.06)	2	1.36	1, 6	0.97	0.08
SC1B1						
ES1	2.04 (.03)	2	1.05	1, 6	1.19	1.50
ES2	1.94 (.05)	2	0.98	1, 6	1.32	2.17
EN	2.29 (.06)	2	1.24	1, 6	0.88	0.22
SC1B3						
ES1	2.63 (.05)	2	1.41	1, 6	0.82	-0.15
ES2	2.28 (.06)	2	1.19	1, 6	0.84	0.21
EN	2.54 (.06)	2	1.36	1, 6	0.75	-0.13
SC1B4						
ES1	1.72 (.03)	1	1.00	1, 6	1.74	3.31
ES2	1.60 (.04)	1	0.93	1, 6	1.98	4.74
EN	2.00 (.06)	1	1.33	1, 6	1.22	0.57
SC1B5						
ES1	1.86 (.04)	1	1.16	1, 6	1.61	2.28
ES2	1.87 (.05)	2	1.10	1, 6	1.45	1.86
EN	2.35 (.07)	2	1.42	1, 6	0.90	-0.03
SC2B1						
ES1	2.03 (.04)	2	1.09	1, 6	1.57	2.78
ES2	1.91 (.05)	2	1.01	1, 6	1.36	2.14
EN	2.48 (.06)	2	1.28	1, 6	0.79	0.05
SC3A2[i]						
ES1	3.45 (.05)	3	1.41	1, 6	0.10	-0.92
ES2	3.81 (.07)	4	1.38	1, 6	-0.11	-0.91
EN	3.48 (.06)	3	1.39	1, 6	0.08	-0.74
SC3A4						
ES1	2.93 (.05)	3	1.52	1, 6	0.59	-0.65

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
ES2	2.54 (.06)	2	1.36	1, 6	0.89	0.07
EN	3.16 (.07)	3	1.54	1, 6	0.04	-0.79
SC3B1						
ES1	4.29 (.04)	4	1.31	1, 6	-0.30	-0.78
ES2	3.67 (.07)	4	1.39	1, 6	0.09	-0.86
EN	3.89 (.07)	4	1.40	1, 6	-0.04	-0.77
SC3B2						
ES1	3.78 (.04)	4	1.33	1, 6	0.07	-0.79
ES2	3.35 (.06)	3	1.28	1, 6	0.43	-0.39
EN	3.60 (.06)	3	1.29	1, 6	0.15	-0.43
SC3B5						
ES1	2.68 (.04)	2	1.39	1, 6	0.77	-0.11
ES2	2.31 (.06)	2	1.21	1, 6	1.13	1.16
EN	2.61 (.07)	2	1.40	1, 6	0.30	-0.11
SC3B6						
ES1	2.40 (.04)	2	1.28	1, 6	1.02	0.54
ES2	2.08 (.05)	2	1.10	1, 6	1.17	1.31
EN	2.51 (.06)	2	1.27	1, 6	0.70	0.07
SC3B7						
ES1	4.08 (.04)	4	1.34	1, 6	-0.19	-0.80
ES2	3.39 (.07)	3	1.40	1, 6	0.27	-0.74
EN	3.79 (.06)	4	1.29	1, 6	0.00	-0.41
SC4A4						
ES1	3.51 (.05)	3	1.42	1, 6	0.19	-0.83
ES2	3.48 (.06)	3	1.34	1, 6	0.20	-0.67
EN	3.86 (.06)	4	1.32	1, 6	0.11	-0.70
SC4A5						
ES1	2.75 (.04)	2	1.37	1, 6	0.73	-0.23
ES2	2.48 (.06)	2	1.28	1, 6	0.85	0.34
EN	3.28 (.07)	3	1.44	1, 6	0.31	-0.73
SC4B3						
ES1	3.63 (.04)	3	1.34	1, 6	0.04	-0.63
ES2	3.51 (.06)	3	1.35	1, 6	0.22	-0.53
EN	3.86 (.06)	4	1.23	1, 6	0.14	-0.51
SC4B4						
ES1	3.23 (.04)	3	1.22	1, 6	0.37	-0.06
ES2	2.72 (.06)	3	1.18	1, 6	0.61	0.24
EN	3.55 (.06)	3	1.25	1, 6		

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
SC4B10						
ES1	3.13 (.04)	3	1.37	1, 6	0.44	-0.48
ES2	3.10 (.06)	3	1.27	1, 6	0.54	-0.26
EN	3.64 (.06)	3	1.34	1, 6	0.15	-0.50
SC4B12						
ES1	3.04 (.05)	3	1.45	1, 6	0.38	-0.70
ES2	2.87 (.06)	3	1.36	1, 6	0.41	-0.49
EN	3.39 (.07)	3	1.49	1, 6	0.16	-0.77
SC4C1[i]						
ES1	4.04 (.04)	4	1.16	1, 6	0.14	-0.44
ES2	4.64 (.05)	5	1.15	1, 6	-0.67	-0.03
EN	3.38 (.06)	4	1.30	1, 6	-0.13	-0.53
SC4D1						
ES1	3.32 (.05)	3	1.60	1, 6	0.28	-1.09
ES2	2.81 (.06)	3	1.43	1, 6	0.59	-0.17
EN	3.23 (.07)	3	1.55	1, 6	0.32	-0.85
SC4D5						
ES1	3.24 (.05)	3	1.48	1, 6	0.35	-0.84
ES2	2.86 (.07)	3	1.39	1, 6	0.47	-0.57
EN	3.31 (.07)	3	1.56	1, 6	0.22	-0.89
SC4D7						
ES1	2.90 (.05)	2	1.56	1, 6	0.62	-0.71
ES2	2.47 (.06)	2	1.34	1, 6	0.93	0.23
EN	2.92 (.07)	3	1.52	1, 6	0.44	-0.71
RA1A2						
ES1	1.96 (.04)	2	1.19	1, 6	1.21	0.83
ES2	1.67 (.05)	1	1.03	1, 6	1.83	3.48
EN	1.86 (.06)	1	1.19	1, 6	1.30	0.90
RA1A3						
ES1	2.99 (.05)	3	1.55	1, 6	0.44	-0.85
ES2	2.69 (.07)	2	1.47	1, 6	0.60	-0.60
EN	2.91 (.07)	3	1.52	1, 6	0.40	-0.80
RA1A5						
ES1	2.16 (.04)	2	1.15	1, 6	1.06	0.76
ES2	1.99 (.05)	2	1.09	1, 6	1.25	1.52
EN	2.50 (.06)	2	1.30	1, 6	0.70	0.01
RA1A6						
ES1	2.25 (.04)	2	1.61	1, 6	0.99	0.39

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
ES2	1.84 (.05)	1.5	1.07	1, 6	1.33	1.37
EN	2.27 (.06)	2	1.28	1, 6	0.91	0.20
RA1A7						
ES1	2.73 (.05)	3	1.48	1, 6	0.51	-0.76
ES2	2.38 (.06)	2	1.36	1, 6	0.83	-0.02
EN	2.61 (.07)	3	1.41	1, 6	0.69	-0.10
RA1A8						
ES1	2.43 (.04)	2	1.32	1, 6	0.77	-0.10
ES2	2.12 (.06)	2	1.29	1, 6	1.20	0.78
EN	2.58 (.06)	2	1.37	1, 6	0.69	-0.16
RA2A2						
ES1	1.86 (.04)	2	1.13	1, 6	1.49	1.89
ES2	1.87 (.05)	2	1.10	1, 6	1.47	1.88
EN	2.58 (.07)	2	1.44	1, 6	0.69	-0.43
RA2A3						
ES1	2.85 (.05)	3	1.40	1, 6	0.06	-0.36
ES2	2.89 (.07)	3	1.40	1, 6	0.44	-0.53
EN	3.38 (.06)	3	1.30	1, 6	0.27	-0.46
RA2A5						
ES1	2.77 (.05)	3	1.46	1, 6	0.61	-0.52
ES2	2.77 (.07)	3	1.50	1, 6	0.61	-0.55
EN	2.98 (.06)	3	1.39	1, 6	0.44	-0.43
RA2A8						
ES1	2.30 (.04)	2	1.18	1, 6	1.01	0.93
ES2	2.29 (.05)	2	1.16	1, 6	0.75	0.19
EN	3.01 (.06)	3	1.25	1, 6	0.48	-0.06
RA2A10						
ES1	2.28 (.04)	2	1.30	1, 6	0.99	0.28
ES2	2.24 (.06)	2	1.28	1, 6	0.93	0.31
EN	2.48 (.06)	2	1.35	1, 6	0.66	-0.34
RA3A1						
ES1	2.41 (.04)	2	1.22	1, 6	0.70	-0.06
ES2	3.32 (.06)	2	1.24	1, 6	0.81	-0.01
EN	2.89 (.06)	3	1.26	1, 6	0.31	-0.28
RA3A2						
ES1	2.76 (.04)	3	1.31	1, 6	0.45	-0.45
ES2	2.45 (.06)	2	1.26	1, 6	0.79	0.14
EN	2.83 (.06)	3	1.22	1, 6	0.39	-0.18

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
RA3A3						
ES1	2.56 (.05)	2	1.48	1, 6	0.73	-0.49
ES2	2.22 (.06)	2	1.33	1, 6	1.02	0.25
EN	2.39 (.06)	2	1.38	1, 6	0.80	-0.21
RA3A4						
ES1	2.09 (.03)	2	1.06	1, 6	1.00	0.96
ES2	1.99 (.05)	2	1.05	1, 6	1.13	1.20
EN	2.40 (.06)	2	1.23	1, 6	0.79	0.34
RA3A6						
ES1	2.21 (.04)	2	1.21	1, 6	0.92	0.29
ES2	1.81 (.05)	2	1.02	1, 6	1.52	2.54
EN	2.21 (.06)	2	1.29	1, 6	0.94	0.22
RA3A8						
ES1	2.18 (.04)	2	1.21	1, 6	0.94	0.23
ES2	1.92 (.05)	2	1.08	1, 6	1.31	1.71
EN	2.39 (.07)	2	1.40	1, 6	0.82	-0.14
RA4A5						
ES1	1.87 (.03)	2	1.02	1, 6	1.48	2.51
ES2	1.81 (.04)	2	0.95	1, 6	1.47	2.71
EN	2.06 (.06)	2	1.32	1, 6	1.20	0.64
RA4B2						
ES1	3.00 (.05)	3	1.40	1, 6	0.27	-0.72
ES2	3.00 (.06)	3	1.34	1, 6	0.32	-0.62
EN	2.98 (.06)	3	1.18	1, 6	0.26	-0.20
RA4B3						
ES1	2.54 (.04)	2	1.11	1, 6	0.68	0.32
ES2	2.52 (.05)	2	1.02	1, 6	0.61	0.41
EN	2.58 (.05)	3	1.10	1, 6	0.61	0.54
RA4B5						
ES1	1.69 (.03)	1	0.98	1, 6	1.64	2.73
ES2	1.71 (.05)	1	1.01	1, 6	1.53	2.20
EN	2.12 (.05)	2	1.14	1, 6	0.82	0.13
RA4B6						
ES1	2.38 (.04)	2	1.15	1, 6	0.60	-0.10
ES2	2.15 (.05)	2	1.16	1, 6	0.93	0.34
EN	2.25 (.06)	2	1.21	1, 6	0.83	0.00
RB1A5						
ES1	3.37 (.04)	3	1.38	1, 6	0.29	-0.79

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
ES2	3.13 (.07)	3	1.45	1, 6	0.46	-0.73
EN	3.12 (.07)	3	1.43	1, 6	0.38	-0.65
RB1A8						
ES1	2.20 (.04)	2	1.09	1, 6	1.18	1.51
ES2	2.20 (.05)	2	1.08	1, 6	1.18	1.57
EN	2.94 (.06)	3	1.30	1, 6	0.61	-0.09
RB1A9						
ES1	2.41 (.04)	2	1.16	1, 6	0.96	0.60
ES2	2.20 (.05)	2	1.08	1, 6	1.20	1.71
EN	2.51 (.06)	2	1.34	1, 6	0.80	0.06
RB1A11[i]						
ES1	3.91 (.04)	4	1.31	1, 6	-0.12	-0.82
ES2	4.26 (.06)	5	1.27	1, 6	-0.60	-0.34
EN	1.73 (.60)	4	1.21	1, 6	-0.20	-0.36
RB1A12[i]						
ES1	3.94 (.04)	4	1.33	1, 6	-0.30	-0.70
ES2	4.24 (.06)	5	1.29	1, 6	-0.50	-0.56
EN	3.63 (.06)	4	1.25	1, 6	-0.16	-0.36
RB2A1						
ES1	2.35 (.04)	2	1.27	1, 6	0.97	0.44
ES2	2.29 (.06)	2	1.26	1, 6	1.22	1.11
EN	2.79 (.07)	3	1.41	1, 6	0.57	-0.42
RB2A2						
ES1	2.62 (.04)	2	1.34	1, 6	0.66	-0.21
ES2	2.45 (.06)	2	1.26	1, 6	0.85	0.25
EN	2.65 (.06)	3	1.37	1, 6	0.50	-0.54
RB2A6						
ES1	2.46 (.05)	2	1.40	1, 6	0.89	0.01
ES2	2.38 (.06)	2	1.25	1, 6	0.87	0.33
EN	2.46 (.06)	2	1.36	1, 6	0.78	-0.09
RB2B4						
ES1	2.79 (.04)	3	1.37	1, 6	0.65	-0.30
ES2	2.57 (.06)	2	1.32	1, 6	0.74	-0.07
EN	2.88 (.07)	3	1.43	1, 6	0.50	-0.52
RB2B6						
ES1	3.87 (.05)	4	1.46	1, 6	-0.14	-0.96
ES2	3.58 (.07)	3	1.40	1, 6	0.20	-0.79
EN	3.56 (.07)	3	1.51	1, 6	0.16	-0.86

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
RB2B7						
ES1	3.76 (.05)	4	1.50	1, 6	-0.04	-1.02
ES2	3.50 (.07)	3	1.46	1, 6	0.24	-0.82
EN	3.42 (.07)	3	1.53	1, 6	0.22	-0.85
RB3A3						
ES1	3.34 (.04)	3	1.23	1, 6	0.32	-0.44
ES2	2.98 (.06)	3	1.19	1, 6	0.54	-0.02
EN	2.73 (.06)	3	1.21	1, 6	0.56	0.03
RB3A5						
ES1	2.36 (.04)	2	1.14	1, 6	1.05	1.07
ES2	2.31 (.05)	2	1.09	1, 6	1.02	1.14
EN	2.64 (.05)	3	1.14	1, 6	0.57	0.15
RB3A8						
ES1	3.19 (.04)	3	1.16	1, 6	0.30	-0.36
ES2	3.20 (.05)	3	1.11	1, 6	0.54	0.14
EN	2.98 (.04)	3	0.94	1, 6	0.27	0.81
RB3B2						
ES1	3.76 (.05)	4	1.51	1, 6	-0.12	-1.07
ES2	3.75 (.07)	4	1.54	1, 6	-0.16	-1.07
EN	3.19 (.07)	3	1.49	1, 6	0.20	-0.75
RB3B3						
ES1	3.12 (.04)	3	1.29	1, 6	0.35	-0.54
ES2	3.34 (.06)	3	1.27	1, 6	0.27	-0.54
EN	3.68 (.05)	4	1.16	1, 6	0.20	-0.18
RB3B12						
ES1	3.37 (.05)	3	1.39	1, 6	0.20	-0.90
ES2	3.43 (.07)	3	1.44	1, 6	0.11	-1.01
EN	3.46 (.06)	3	1.33	1, 6	0.12	-0.50
RB3B13						
ES1	2.55 (.04)	2	1.21	1, 6	0.78	0.15
ES2	2.53 (.06)	2	1.30	1, 6	0.78	-0.04
EN	3.11 (.06)	3	1.31	1, 6	0.32	-0.51
RB4A1						
ES1	2.11 (.04)	2	1.09	1, 6	0.82	0.11
ES2	2.18 (.05)	2	1.12	1, 6	0.82	0.26
EN	2.48 (.06)	2	1.22	1, 6	0.67	0.12
RB4A2						
ES1	3.06 (.05)	3	1.49	1, 6	0.27	-0.95

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
ES2	3.27 (.07)	3	1.57	1, 6	0.10	-1.12
EN	2.71 (.06)	3	1.22	1, 6	0.44	-0.17
RB4A6						
ES1	2.45 (.04)	2	1.35	1, 6	0.94	0.21
ES2	2.15 (.06)	2	1.30	1, 6	1.24	1.02
EN	2.94 (.07)	3	2.30	1, 6	0.42	-0.72
RB4A13						
ES1	2.17 (.04)	2	1.29	1, 6	1.03	0.25
ES2	2.16 (.06)	2	1.25	1, 6	1.10	0.57
EN	3.08 (.06)	3	1.23	1, 6	0.19	-0.24
RB4A15						
ES1	2.27 (.04)	2	1.26	1, 6	0.91	0.16
ES2	2.26 (.06)	2	1.27	1, 6	0.96	0.29
EN	2.86 (.06)	3	1.23	1, 6	0.30	-0.38
RB5A2						
ES1	3.29 (.04)	3	1.32	1, 6	-0.06	-0.73
ES2	2.80 (.06)	3	1.34	1, 6	0.34	-0.71
EN	3.25 (.06)	3	1.39	1, 6	0.15	-0.58
RB5A3						
ES1	1.71 (.03)	1	1.02	1, 6	1.90	3.97
ES2	1.82 (.05)	2	1.06	1, 6	1.68	3.35
EN	2.41 (.07)	2	1.42	1, 6	0.81	-0.27
RB5A4						
ES1	2.08 (.04)	2	1.20	1, 6	1.35	1.58
ES2	2.01 (.05)	2	0.16	1, 6	1.28	1.38
EN	2.42 (.07)	2	1.44	1, 6	0.89	-0.13
RB5A9						
ES1	2.53 (.04)	2	1.29	1, 6	0.76	-0.03
ES2	2.44 (.06)	2	1.34	1, 6	0.98	0.26
EN	2.80 (.06)	3	1.24	1, 6	0.06	0.16
RB5B8						
ES1	1.61 (.03)	1	0.96	1, 6	2.10	4.99
ES2	1.55 (.04)	1	0.09	1, 6	2.08	4.95
EN	196 (.06)	1	1.34	1, 6	1.33	0.89
RC1A2						
ES1	3.40 (.04)	3	1.31	1, 6	0.29	-0.75
ES2	2.93 (.06)	3	1.22	1, 6	0.80	0.33
EN	3.18 (.06)	3	1.36	1, 6	0.56	-0.33

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
RC1A4						
ES1	3.76 (.05)	4	1.43	1, 6	0.09	-0.99
ES2	3.59 (.06)	3	1.30	1, 6	0.23	-0.62
EN	3.51 (.06)	3	1.38	1, 6	0.28	-0.61
RC1A5						
ES1	3.06 (.04)	3	1.29	1, 6	0.57	-0.27
ES2	2.84 (.05)	3	1.16	1, 6	0.87	0.49
EN	2.62 (.06)	2	1.75	1, 6	0.83	0.18
RC1A7						
ES1	3.48 (.04)	3	1.38	1, 6	0.16	-0.81
ES2	3.13 (.06)	3	1.32	1, 6	0.51	-0.36
EN	3.08 (.06)	3	1.29	1, 6	0.52	-0.13
RC1A9						
ES1	2.83 (.04)	3	1.36	1, 6	-0.65	-0.28
ES2	2.55 (.06)	2	1.18	1, 6	0.81	0.43
EN	3.03 (.06)	3	1.37	1, 6	0.51	-0.40
RC1A11						
ES1	3.99 (.04)	4	1.35	1, 6	-0.11	-0.86
ES2	3.45 (.06)	3	1.28	1, 6	0.29	-0.55
EN	3.46 (.06)	3	1.33	1, 6	0.40	-0.43
RC2A3						
ES1	3.20 (.05)	3	1.47	1, 6	0.36	-0.76
ES2	2.77 (.07)	3	1.43	1, 6	0.51	-0.48
EN	2.84 (.06)	3	1.27	1, 6	0.46	-0.12
RC2A4						
ES1	2.69 (.04)	3	1.33	1, 6	0.73	0.01
ES2	2.48 (.06)	2	1.28	1, 6	0.79	0.19
EN	3.24 (.06)	3	1.30	1, 6	0.20	-0.41
RC2A6						
ES1	3.18 (.05)	3	1.55	1, 6	0.23	-0.99
ES2	2.97 (.07)	3	1.52	1, 6	0.40	-0.85
EN	3.41 (.07)	3	1.42	1, 6	0.08	-0.65
RC2B5						
ES1	2.62 (.05)	2	1.48	1, 6	0.74	-0.39
ES2	2.46 (.07)	2	1.48	1, 6	0.09	-0.18
EN	2.57 (.07)	2	1.41	1, 6	0.61	-0.41
RC2C2						
ES1	3.49 (.05)	3	1.46	1, 6	0.12	-0.95

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
ES2	3.18 (.07)	3	1.41	1, 6	0.35	-0.66
EN	3.03 (.06)	3	1.36	1, 6	0.42	-0.42
RC2C3						
ES1	3.32 (.05)	3	1.41	1, 6	0.28	-0.86
ES2	2.98 (.06)	3	1.37	1, 6	0.44	-0.52
EN	2.86 (.06)	3	1.37	1, 6	0.51	-0.28
RC2C4						
ES1	3.71 (.05)	4	1.44	1, 6	0.01	-0.97
ES2	3.43 (.07)	3	1.51	1, 6	0.15	-0.95
EN	3.17 (.06)	3	1.30	1, 6	0.34	-0.32
RC2C5						
ES1	3.43 (.05)	3	1.46	1, 6	0.18	-0.90
ES2	3.54 (.07)	3	1.50	1, 6	0.14	-0.94
EN	3.02 (.06)	3	1.31	1, 6	0.44	-0.13
RC2C7						
ES1	3.18 (.05)	3	1.47	1, 6	0.36	-0.88
ES2	3.02 (.06)	3	1.37	1, 6	0.50	-0.41
EN	3.13 (.07)	3	1.40	1, 6	0.34	-0.58
RD1A2						
ES1	2.24 (.04)	2	1.26	1, 6	0.97	0.24
ES2	2.23 (.07)	2	1.40	1, 6	1.03	0.09
EN	2.44 (.07)	2	1.40	1, 6	0.75	-0.22
RD1A3						
ES1	1.49 (.03)	1	0.87	1, 6	2.50	7.46
ES2	1.52 (.04)	1	0.85	1, 6	2.17	6.03
EN	1.87 (.05)	1	1.70	1, 6	1.47	1.67
RD1A5						
ES1	2.37 (.04)	2	1.28	1, 6	0.82	-0.06
ES2	2.19 (.06)	2	1.30	1, 6	1.07	0.46
EN	2.48 (.06)	2	1.35	1, 6	0.66	-0.37
RD1A10						
ES1	1.87 (.04)	2	1.12	1, 6	1.55	2.30
ES2	1.75 (.05)	1	1.11	1, 6	1.84	3.35
EN	2.12 (.06)	2	1.24	1, 6	0.98	0.22
RD2A1						
ES1	3.24 (.05)	3	1.39	1, 6	0.42	-0.06
ES2	3.05 (.07)	3	1.45	1, 6	0.50	-0.56
EN	3.17 (.06)	3	1.38	1, 6	0.35	-0.49

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
RD2A2						
ES1	3.64 (.05)	3	1.50	1, 6	0.10	-1.00
ES2	3.39 (.07)	3	1.49	1, 6	0.34	-0.75
EN	3.20 (.07)	3	1.49	1, 6	0.35	-0.73
RD2A5						
ES1	3.33 (.05)	3	1.55	1, 6	0.22	-1.00
ES2	3.08 (.07)	3	1.53	1, 6	0.42	-0.82
EN	2.68 (.07)	2	1.51	1, 6	0.77	-0.30
RD2A6						
ES1	3.06 (.05)	3	1.60	1, 6	0.39	-0.98
ES2	2.85 (.07)	3	1.47	1, 6	0.59	-0.53
EN	3.02 (.07)	3	1.58	1, 6	0.45	-0.78
RD2A8						
ES1	2.86 (.05)	3	1.50	1, 6	0.50	-0.75
ES2	2.80 (.07)	3	1.44	1, 6	0.64	-0.46
EN	2.74 (.07)	3	1.48	1, 6	0.56	-0.57
RD2A10						
ES1	2.96 (.05)	3	1.55	1, 6	0.53	-0.79
ES2	2.74 (.07)	2	1.45	1, 6	0.69	-0.31
EN	2.90 (.07)	3	1.57	1, 6	0.49	-0.78
RD3A1						
ES1	3.00 (.05)	3	1.49	1, 6	0.42	-0.76
ES2	2.95 (.07)	3	1.45	1, 6	0.50	-0.61
EN	2.90 (.07)	3	1.43	1, 6	0.54	-0.37
RD3A3						
ES1	2.61 (.05)	2	1.42	1, 6	0.72	-0.31
ES2	2.32 (.06)	2	1.36	1, 6	0.96	0.23
EN	2.43 (.06)	2	1.33	1, 6	0.69	-0.28
RD3A4						
ES1	1.78 (.03)	1	1.04	1, 6	1.50	2.09
ES2	1.52 (.04)	1	0.89	1, 6	1.98	4.29
EN	1.85 (.06)	1	1.23	1, 6	1.36	0.84
RD3A7						
ES1	2.19 (.04)	2	1.21	1, 6	1.00	0.52
ES2	2.18 (.06)	2	1.24	1, 6	1.03	0.54
EN	2.38 (.06)	2	1.33	1, 6	0.83	0.08
RD3A10						
ES1	3.30 (.05)	3	1.52	1, 6	0.21	-0.90

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
ES2	3.28 (.07)	3	1.51	1, 6	0.23	-0.94
EN	2.93 (.06)	3	1.37	1, 6	0.44	-0.35
RD3A12						
ES1	1.79 (.04)	1	1.11	1, 6	1.67	2.70
EN	2.30 (.06)	2	1.35	1, 6	0.91	0.13
RD3A14						
ES1	2.11 (.04)	2	1.32	1, 6	1.21	0.76
ES2	2.17 (.06)	2	1.32	1, 6	1.20	0.88
EN	2.50 (.06)	2	1.39	1, 6	0.70	-0.25
RD4A2						
ES1	2.88 (.04)	3	1.34	1, 6	0.49	-0.38
ES2	2.52 (.06)	2	1.24	1, 6	0.52	-0.37
EN	3.09 (.06)	3	1.33	1, 6	0.21	-0.38
RD4A3						
ES1	2.91 (.05)	3	1.44	1, 6	0.47	-0.59
ES2	0.52 (.07)	2	1.38	1, 6	0.75	-0.11
EN	2.64 (.06)	3	1.31	1, 6	0.42	-0.51
RD4A4						
ES1	2.52 (.04)	2	1.27	1, 6	0.69	0.01
ES2	2.45 (.06)	2	1.28	1, 6	0.82	0.33
EN	3.36 (.06)	3	1.32	1, 6	0.25	-0.28
RD4A6						
ES1	2.62 (.04)	2	1.36	1, 6	0.66	-0.27
ES2	2.20 (.06)	2	1.22	1, 6	0.96	0.47
EN	2.59 (.06)	3	1.35	1, 6	0.51	-0.41
RD4A7						
ES1	2.18 (.04)	2	1.27	1, 6	1.05	0.47
ES2	1.92 (.05)	2	1.14	1, 6	1.38	1.60
EN	2.26 (.06)	2	1.30	1, 6	0.09	0.01
Alternative Pool of Items						
AQ39[i]						
ES1	2.32 (.04)	2	1.20	1, 6	0.98	0.62
ES2	2.06 (.05)	2	1.04	1, 6	1.10	1.47
EN	2.60 (.06)	2	1.31	1, 6	0.62	-0.19
AQ44						
ES1	3.80 (.05)	4	1.39	1, 6	-0.09	-0.79
ES2	4.15 (.06)	4	1.31	1, 6	-0.26	-0.80
EN	3.56 (.06)	3	1.38	1, 6	-0.01	-0.67

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
AQ47						
ES1	3.82 (.05)	4	1.44	1, 6	-0.09	-0.92
ES2	4.20 (.06)	4	1.33	1, 6	-0.27	-0.09
EN	3.58 (.06)	3	1.37	1, 6	0.13	-0.68
BAPQ1						
ES1	4.14 (.04)	4	1.25	1, 6	-0.22	-0.64
ES2	4.41 (.06)	5	1.20	1, 6	-0.29	-0.83
EN	3.68 (.06)	4	1.31	1, 6	-0.03	-0.63
BAPQ17						
ES1	2.44 (.04)	2	1.32	1, 6	0.85	0.10
ES2	2.13 (.05)	2	1.10	1, 6	1.07	1.15
EN	2.76 (.05)	3	1.23	1, 6	0.33	0.13
BAPQ36[i]						
ES1	4.09 (.04)	4	1.31	1, 6	-0.22	-0.74
ES2	4.37 (.06)	5	1.25	1, 6	-0.37	-0.82
EN	3.76 (.06)	4	1.31	1, 6	0.00	-0.56
SRS12[i]						
ES1	4.11 (.04)	4	1.36	1, 6	-0.37	-0.75
ES2	4.32 (.06)	5	1.35	1, 6	-0.47	-0.74
EN	3.92 (.05)	4	1.22	1, 6	-0.12	-0.52
SRS31						
ES1	3.45 (.04)	3	1.30	1, 6	0.10	-0.51
ES2	3.13 (.06)	3	1.31	1, 6	0.26	-0.44
EN	3.37 (.06)	3	1.35	1, 6	0.17	-0.57
SRS60						
ES1	2.49 (.05)	2	1.42	1, 6	0.80	-0.22
ES2	2.21 (.06)	2	1.31	1, 6	1.01	0.32
EN	2.61 (.06)	3	1.38	1, 6	0.57	-0.37
SRS64						
ES1	3.29 (.05)	3	1.58	1, 6	0.31	-0.98
ES2	2.79 (.07)	3	1.43	1, 6	0.55	-0.54
EN	3.34 (.07)	3	1.52	1, 6	0.19	-0.85
			BAPQ Subscales			
Aloof						
ES1	3.09 (.03)	3.00	.91	1.25, 5.67	0.47	-0.27
ES2	2.78 (.04)	2.67	.93	1.00, 5.92	0.51	-0.08
EN	3.24 (.04)	3.25	.94	1.08, 5.75	0.04	-0.28
Prag Lang						

Item code	M (SE)	Mdn	SD	Min, Max	Sk	K
ES1	2.57 (.03)	2.50	.79	1.00, 5.92	0.66	0.62
ES2	2.52 (.03)	2.50	.59	1.08, 4.5	0.40	0.26
EN	2.85 (.03)	2.83	.78	1.00, 5.17	0.11	-0.27
Rigid						
ES1	2.99 (.03)	2.92	.83	1.00, 5.75	0.47	0.27
ES2	2.84 (.04)	2.75	.83	1.17, 5.75	0.54	0.49
EN	3.24 (.04)	2.25	.85	1.00, 6.00	0.09	0.26

Note. M: mean; SE: standard error; Mdn: median; SD: standard deviation; Min, Max: minimum, maximum; Sk: skewness; K: kurtosis.

[i] Inverted items.

Table 5
Mardia's Multivariate Skewness and Kurtosis

SOCIAL-BAP	*BAP-IT Candidate Items*		*Alternative Pool of Items*	
Sample	**SK**	**K**	**SK**	**K**
ES1	123377.38	177.43	161302.32	705.53
ES2	114391.07	117.98	3076.68	52.42
EN	119057.66	145.68	60820.19	345.91
RIRE-BAP	*BAP-IT Candidate Items*		*Alternative Pool of Items*	
Sample	**SK**	**K**	**SK**	**K**
ES1	199012.29	122.42	158572.06	733.09
ES2	184705.01	107.32	3076.68	52.42
EN	190590.24	122.43	60820.19	345.91
TOTAL	*BAP-IT Candidate Items*		*Alternative Pool of Items*	
Sample	**SK**	**K**	**SK**	**K**
ES1	1029235.92	254.75	176132.11	419.37
ES2	892427.49	113.78	11911.15	71.04
EN	936092.75	140.01	157358.51	476.37

Note. Sk: skewness; K: kurtosis.
All values were statistically significant. Significant level was fixed at $p < .001$ using Bonferroni correction.

Table 6

SOCIAL-BAP Unrotated Factor Matrix from Step 1: Exploratory Item Inspection According to Unidimensionality

	Factor											
Item code	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
SA1A1	**.39**[a]	.06	.15	-.10	.06	.03	-.03	-.08	.04	.11	-.01	-.11
SA1A2	**.38**[a]	-.02	.12	-.17	.10	.08	-.07	-.03	-.02	.24	-.04	-.11
SA1A3	**.58**[a]	.05	.04	-.23	.10	.01	.02	-.18	.08	.16	.02	.02
SA1A7	**.69**[a]	.11	-.15	-.26	.02	.02	-.08	-.22	.06	-.13	.09	-.19
SA1B2	**.52**[a]	-.11	.23	.11	.20	-.04	.04	.08	-.08	-.07	.16	-.13
SA1B3	**.34**	-.17	**.35**	.23	.19	-.09	.09	.26	-.06	-.08	.13	-.17
SA1B4	.27	-.16	**.45**	.13	.20	-.09	.05	.18	.05	.02	.10	-.08
SA1B10[i]	**-.40**[a]	.19	-.15	-.20	.05	.21	.06	-.12	.11	.04	-.03	.15
SA2A2	**.61**[a]	.25	-.08	.12	-.06	.04	-.03	-.06	.03	.11	.04	.02
SA2A3	**.57**	**.41**	-.09	.19	.05	.07	-.04	-.02	-.05	-.04	-.02	.02
SA2A5	**.65**[a]	.27	-.08	.24	-.01	-.02	.01	.00	-.02	.19	.03	-.10
SA2A10i	**-.67**	**-.37**	.17	-.16	.07	.04	.02	.04	.09	-.07	.03	.12
SA2A17	**.65**	**.39**	-.09	**.33**	-.04	.05	-.03	-.03	-.02	.00	-.03	.01
SA3A2i	**-.48**[a]	-.15	.06	.21	.01	.06	.20	-.11	.14	-.13	.10	.11
SA3A3	**.61**[a]	.06	.06	-.18	-.02	-.08	.07	-.03	.14	.07	.03	.04
SA3A7	**.56**[a]	.16	-.14	-.21	-.02	.00	.22	-.07	.07	.04	.06	.04
SA3A8	**.62**[a]	.26	.13	-.09	.05	-.09	.03	.10	.03	.12	-.04	.11
SA3A9	**.57**[a]	.13	-.01	-.18	.10	.09	-.15	.11	-.08	.07	-.10	.01
SA3A11	**.50**[a]	.21	-.08	-.12	-.07	-.06	.15	.00	.10	.12	.06	-.03
SA3A13	**.69**[a]	.14	-.13	-.20	-.08	.05	-.02	-.15	.12	-.05	.15	-.21
SA3A14	**.68**[a]	.23	.08	-.03	-.03	-.06	.00	.04	.10	.13	-.02	.12
SA3A16[i]	**-.55**[a]	-.06	-.01	.16	.15	.18	.01	.01	-.02	.00	.02	-.03
SB1A2	**.47**[a]	-.28	.13	.13	.18	-.07	.10	-.29	.02	.01	.04	-.02
SB1A7[i]	**-.46**	**.35**	-.07	-.12	-.04	.24	.03	.29	.13	.08	.08	.02
SB1B2	**.61**	**-.34**	-.01	.07	.13	.02	-.06	.09	.00	-.03	-.06	-.07
SB1B4[i]	**-.57**	**.46**	.12	-.07	.07	.13	.13	-.02	.10	.02	.03	.07
SB1C6	**.67**[a]	-.25	-.17	-.04	.12	.14	-.11	.07	.04	-.13	.13	.07
SB1C7	**.64**	**-.31**	-.12	-.01	.13	.05	-.13	.08	.05	-.15	.06	.11
SB1C9	**.57**	**-.34**	-.06	.06	.11	.08	-.05	.06	.02	-.01	-.09	-.09
SB2A1	**.48**	-.07	**.36**	-.01	-.26	.29	.03	-.03	-.01	.06	.00	-.03
SB2A2	**.54**	-.05	**.41**	-.07	-.31	**.30**	.04	-.04	-.05	.00	.03	.04
SB2A4	**.65**[a]	-.21	.29	.01	-.16	.15	-.04	.04	.11	.00	.01	-.02
SB2A6	**.62**	-.07	**.42**	-.11	-.25	.21	.02	.00	-.13	-.06	-.02	.07
SB2A7[i]	**-.53**	.26	**-.33**	.06	**.37**	.04	.08	.00	.12	.00	.03	.03
SB2B1[i]	**-.54**	**.30**	.10	.05	**.35**	**.36**	-.02	-.04	-.11	.02	.04	-.11

	Factor											
Item code	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
SB2B2[i]	**-.49**	.28	.16	-.01	**.33**	.28	-.05	-.05	.00	.07	-.01	-.07
SB2B5	**.55**	**-.30**	-.01	.00	-.06	-.15	.08	.04	.20	.00	-.04	.11
SB3A6	**.37**	-.18	**-.35**	-.05	-.03	.12	.18	.05	-.20	.09	.15	.07
SB3B1	**.54**	**-.39**	.03	.07	.17	.00	.13	-.18	-.06	.04	-.19	.04
SB3B2	**.54**	**-.32**	.11	.13	.19	-.03	.16	-.29	-.03	-.11	-.06	.04
SB3B3	**.48**	**-.41**	-.01	.16	.15	-.03	.19	-.14	-.10	.08	-.09	.06
SB4A1	**.47**	**-.35**	**-.34**	-.06	-.05	.14	.12	.03	-.17	.02	-.01	.02
SB4A9	**.38**	**-.31**	**-.38**	-.03	-.07	.15	.28	.04	-.35	.07	.09	.03
SB4B1	**.67**	-.16	**.42**	-.06	-.22	.21	-.07	-.03	-.05	-.08	.00	.03
SB4C2	**.48**[a]	-.27	-.25	-.05	-.03	.13	.11	.06	.02	.11	.18	-.06
SB4C3	**.55**	**-.33**	-.25	-.02	-.01	.15	-.02	.09	.15	.03	.00	.03
SB4C6	**.59**	**-.40**	-.27	-.04	.06	.12	-.08	.04	.06	-.04	-.10	-.01
SB4C7	**.48**[a]	-.12	-.18	-.04	.03	.04	-.02	.15	.08	-.11	.22	.13
SC1A2	**.59**[a]	-.11	.09	.12	.00	-.07	.12	-.03	.20	.10	.16	-.07
SC1A3	**.66**[a]	-.09	.08	.08	.07	-.11	.04	.04	.07	.24	-.02	.00
SC1B1	**.59**	**-.37**	-.11	-.03	.13	.17	-.11	.11	.09	-.03	-.12	-.03
SC1B3	**.41**[a]	-.16	.26	.13	.14	-.08	.13	-.02	.06	-.04	.06	.09
SC1B4	**.55**	**-.39**	-.15	.01	.08	.12	-.18	.08	.13	.07	-.16	.01
SC1B5	**.47**[a]	-.20	-.09	-.04	.09	.11	-.20	.00	.07	.04	-.06	-.06
SC2B1	**.51**[a]	-.21	.01	.01	-.09	-.09	.05	.03	.14	-.08	-.07	.02
SC3A2[i]	**-.66**[a]	-.25	.20	.19	.10	.13	.09	.11	.04	.16	.01	.13
SC3A4	**.74**[a]	.21	-.11	-.19	.02	-.10	-.09	-.13	-.02	-.15	.08	-.05
SC3B1	**.49**	**.38**	.01	-.24	.09	.08	**.39**	.13	.03	-.18	-.22	-.08
SC3B2	**.57**	**.30**	.06	.07	.05	.01	.09	.00	.06	-.01	-.03	-.01
SC3B5	**.61**[a]	.20	.02	-.04	.04	-.19	.10	.19	-.09	.02	.03	.03
SC3B6	**.74**[a]	.14	-.05	.03	-.03	-.15	.06	.10	.01	.04	.00	-.01
SC3B7	**.55**	**.37**	.05	-.22	.08	.01	**.36**	.17	.01	-.16	-.20	-.09
SC4A4	**.66**	**.33**	.06	.19	.00	.08	-.03	-.02	.05	-.06	-.01	.09
SC4A5	**.78**[a]	.28	-.05	.08	.05	-.07	-.09	.06	-.12	.07	-.06	.02
SC4B3	**.45**	**.42**	-.06	**.34**	.00	.14	-.03	.03	.09	-.08	-.06	.07
SC4B4	**.60**	**.38**	-.17	**.34**	-.02	.16	-.04	-.02	.02	-.04	-.02	.04
SC4B10	**.67**	**.35**	-.02	.29	-.02	.09	.01	-.09	-.02	-.01	.02	.07
SC4B12	**.69**	**.35**	-.15	.26	-.09	.09	-.06	-.03	-.01	-.08	.05	-.01
SC4C1[i]	**-.54**[a]	-.05	.11	-.12	.23	.28	.15	-.06	.11	-.01	.08	.01
SC4d1	**.72**[a]	.26	.12	-.26	.19	-.08	-.09	-.02	-.12	-.05	.05	.09
SC4d5	**.50**	.19	.22	-.27	**.31**	-.04	-.16	.01	-.15	.01	.04	.19
SC4d7	**.70**[a]	.18	.14	-.17	.24	-.10	-.17	-.03	-.10	-.03	.09	.17

Note. Items with loadings above .30 in the primary factor and without cross-loadings on secondary factors were selected.
Loadings higher than .30 are marked in Bold.
[i] Inverted items.
[a] Selected items.

Item code	Factor 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
RA1A2	**.51**[a]	.25	.01	.10	-.09	-.08	.03	-.10	-.03	.05	.03	.11	-.07	.12	.03	-.03	-.06	.01
RA1A3	**.60**[a]	.10	.05	.13	-.05	-.05	.02	-.27	-.16	.12	-.02	.16	.01	.06	-.08	.00	-.12	.11
RA1A5	**.49**[a]	.14	.13	.06	-.11	-.12	-.05	-.08	.02	-.07	.00	.10	-.01	.01	.12	.09	.04	-.06
RA1A6	**.61**[a]	.19	.10	.19	-.09	-.09	-.01	-.27	-.10	.10	.02	.21	.04	.10	.04	.04	-.03	.05
RA1A7	**.66**[a]	.11	.03	.18	-.12	-.02	-.03	-.21	-.15	.05	.00	.21	-.11	.14	.02	-.02	-.01	.06
RA1A8	**.62**[a]	.16	.00	.09	-.20	-.13	.01	-.11	-.03	-.03	.01	.16	-.03	.06	.06	.04	.05	-.08
RA2A2	**.40**[a]	.01	-.17	-.05	.09	-.14	-.22	-.05	.14	.05	.06	.02	.07	.00	.01	.05	-.19	-.01
RA2A3	**.47**	.01	**-.40**	.03	.14	-.17	-.42	.09	.12	.18	.06	-.03	-.05	.02	-.02	.02	-.03	-.06
RA2A5	**.49**	.10	**-.37**	.02	.13	-.13	-.39	.10	.15	.14	-.01	.01	.03	.09	.04	-.04	.05	-.01
RA2A8	**.38**	.03	**-.32**	-.04	.15	-.18	-.45	.05	.29	.16	.06	-.01	.07	-.02	.00	.02	-.10	-.02
RA2A10	**.53**[a]	.15	-.17	.03	-.03	-.12	-.16	.01	.03	.08	.00	.06	-.01	.07	.02	-.02	.05	-.06
RA3A1	**.56**	**.31**	**.33**	.10	-.04	.22	-.19	-.03	-.08	.07	.05	.09	-.11	-.06	-.03	-.01	.09	-.20
RA3A2	**.48**	.25	**.32**	.13	-.02	.22	-.20	-.09	-.13	.14	.03	.10	-.10	-.01	-.07	-.06	.12	-.22
RA3A3	**.53**[a]	.06	.07	.06	-.05	.03	.01	-.01	.03	-.02	.05	-.01	.03	.03	-.08	.09	-.01	.05
RA3A4	**.48**[a]	.16	.12	.00	-.11	-.08	.03	-.03	.05	-.07	.07	.04	-.01	.01	.04	.07	.04	-.01
RA3A6	**.60**[a]	.25	.24	.05	.01	.05	.00	.00	.14	-.15	.08	-.03	-.03	.06	-.01	.03	-.11	.19
RA3A8	**.56**	**.33**	.21	.08	.02	.07	-.01	.04	.13	-.19	.17	-.10	-.09	.04	-.05	-.06	-.06	.11
RA4A5	**.39**[a]	.14	.17	.13	-.05	-.06	.01	.01	.13	-.05	-.04	.01	.02	.04	-.05	.17	-.07	-.07
RA4B2	**.53**[a]	.29	.14	.09	.06	.25	-.09	.06	.04	.01	.01	-.05	-.02	.02	-.01	.03	.08	.01
RA4B3	**.57**[a]	.16	.12	.02	.01	.08	-.07	.05	.06	-.07	-.01	.03	.02	-.04	.06	.14	.03	.09
RA4B5	**.41**	**.31**	.16	-.01	-.01	-.04	-.02	.05	.17	-.04	.03	-.05	.08	-.09	.15	.00	-.07	-.11

	Factor																	
Item code	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
RA4B6	**.53** [a]	.29	.24	.07	-.06	.14	-.03	.05	.10	-.13	.19	-.13	-.03	.05	.05	.01	-.01	.03
RB1A5	**.50**	**-.39**	.06	-.13	.06	-.10	-.05	-.10	-.01	-.03	-.09	-.09	-.05	.12	.04	-.04	.09	.03
RB1A8	**.52** [a]	-.16	.08	-.27	-.01	-.06	-.02	.04	-.01	-.11	.11	-.03	.01	.00	.11	.05	-.12	-.05
RB1A9	**.56** [a]	-.21	.10	-.08	.02	-.02	.03	.03	-.02	.03	.04	-.01	-.01	.01	.04	-.09	-.09	.02
RB1A11[i]	**-.38**	**.37**	-.14	.18	.11	**.32**	.14	.28	-.06	.19	.27	.14	.14	.11	.12	.06	.02	.05
RB1A12[i]	**-.43**	**.42**	-.10	.18	.08	**.32**	.17	.28	-.04	.19	.24	.14	.17	.09	.09	.14	-.02	.05
RB2A1	**.59** [a]	-.25	.15	-.12	.02	-.05	-.05	-.13	-.08	.00	.12	-.10	.02	-.01	.10	-.15	-.03	.00
RB2A2	**.56** [a]	-.13	.12	-.11	.11	.08	.14	.03	-.10	-.01	-.01	-.22	.12	.27	-.02	.08	-.02	-.18
RB2A6	**.46**	-.07	.13	-.11	.12	.05	.05	.04	-.05	-.06	-.04	-.21	.08	**.44**	.00	.09	-.03	-.16
RB2B4	**.54**	**-.36**	.10	-.03	.01	-.11	-.05	-.07	-.09	-.10	.12	-.02	.10	-.12	-.07	.12	.01	.05
RB2B6	**.50**	**-.55**	-.02	.05	-.05	-.02	-.15	-.13	-.19	-.08	**.32**	-.06	.11	-.08	-.13	.10	.07	.08
RB2B7	**.51**	**-.51**	-.02	.06	-.07	.00	-.13	-.10	-.18	-.07	**.33**	-.05	.12	-.03	-.16	.10	.12	.02
RB3A3	**.33** [a]	-.11	.10	.15	-.01	-.01	.05	.07	.01	.08	-.23	.00	-.02	.15	.08	.07	-.01	.15
RB3A5	**.32** [a]	-.06	.20	.12	.02	-.02	-.01	.00	.08	.02	-.07	.00	.05	-.04	.20	.03	-.01	.07
RB3A8	**.34** [a]	-.14	.09	.04	.05	-.01	-.07	.07	.00	-.02	-.13	.08	-.04	.08	.08	-.07	.17	.19
RB3B2	**.38** [a]	-.02	-.12	-.05	.24	.13	-.13	.00	-.07	.08	.04	-.02	-.04	.06	-.04	-.18	.12	.04
RB3B3	**.46** [a]	-.18	.15	-.01	.14	-.02	-.01	-.06	-.10	.05	-.07	.05	.11	-.06	.18	-.06	-.01	.13
RB3B12	**.41** [a]	-.19	.13	-.14	.10	-.03	.00	-.06	-.16	.10	.11	-.01	.17	-.14	.29	-.18	-.03	-.03
RB3B13	**.46** [a]	-.08	.20	-.24	.06	-.08	.06	.05	-.12	-.02	.02	-.03	.20	-.03	.25	-.04	-.02	-.09
RB4A1	**.48**	**.36**	**.31**	.01	.06	.12	-.09	.09	.16	-.16	.16	-.13	-.03	-.04	-.09	-.07	-.04	.12
RB4A2	**.37** [a]	.17	.22	.01	.16	.17	-.13	.05	-.12	.08	-.19	-.06	.01	-.08	-.02	.02	.00	.11

	Factor																	
Item code	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
RB4A6	**.60** [a]	-.06	.19	-.09	.02	.02	-.07	.01	-.07	-.06	-.14	-.03	-.04	.01	.05	.05	.02	.07
RB4A13	**.46**	.23	**.34**	.01	.09	.21	-.21	.04	-.17	.07	-.19	-.08	.03	-.14	-.05	-.02	-.05	.01
RB4A15	**.49**	**.31**	**.35**	.05	.08	.25	-.20	.06	-.13	.06	-.22	-.12	.08	-.17	-.12	.03	-.08	-.05
RB5A2	**.37** [a]	-.10	.20	-.21	.15	-.05	.03	.16	-.10	.23	.07	.13	-.04	.11	-.07	-.14	.09	.01
RB5A3	**.40**	-.02	.23	-.32	.11	-.15	.15	.15	.02	.09	-.01	.03	-.11	.07	-.13	.05	-.03	.02
RB5A4	**.30**	-.12	**.34**	-.42	.17	-.23	.15	.20	.03	.18	.06	.24	-.07	-.06	-.12	.03	-.02	.00
RB5A9	**.54** [a]	-.12	.19	-.23	.04	-.14	.09	.11	.05	.09	.01	-.05	-.12	.16	-.02	.01	.10	.01
RB5B8	**.33**	-.04	**.31**	-.37	.10	-.17	.17	.22	.07	.13	.08	.14	-.12	-.18	-.10	.07	-.04	-.02
RC1A2	**.38**	**-.52**	.12	**.41**	-.06	-.04	.05	.24	.08	.03	-.09	.02	.03	-.05	-.02	-.05	-.01	-.08
RC1A4	**.43**	**-.51**	.00	**.38**	-.06	-.04	.03	.18	.01	.02	.01	.10	-.03	-.02	-.06	-.05	-.05	-.04
RC1A5	**.43**	**-.45**	.17	**.35**	-.04	-.03	.07	.20	.12	.02	-.02	-.04	.04	-.05	.03	-.04	-.04	-.04
RC1A7	**.50**	**-.50**	.09	.28	-.06	.01	.03	.15	.03	.00	-.06	.02	-.05	.08	.00	-.02	.00	.01
RC1A9	**.50**	**-.43**	.13	**.30**	-.11	-.08	.04	.18	.11	-.07	-.07	.01	.00	.01	.00	.02	.04	.02
RC1A11	**.34**	**-.53**	.01	**.39**	-.06	.04	.02	.19	.01	.04	-.01	.04	.06	-.05	-.01	-.03	-.06	-.06
RC2A3	**.40**	.23	**-.34**	-.05	-.01	-.04	-.02	.24	-.28	-.13	-.09	-.04	.03	-.01	-.01	.04	-.02	.01
RC2A4	**.40** [a]	.08	-.27	-.13	-.04	-.19	-.11	.23	-.27	-.11	-.10	.06	.06	-.04	-.02	.18	.07	.09
RC2A6	**.40**	.07	**-.32**	-.07	.00	-.09	-.11	.16	-.14	-.09	-.09	-.01	.10	.01	.01	.08	.12	.05
RC2B5	**.45** [a]	.20	-.18	-.01	.00	-.09	-.07	.10	.02	-.08	.00	.02	-.14	.05	.09	-.03	.08	.11
RC2C2	**.51**	-.01	-.26	.06	**.41**	.13	.15	.00	.01	-.13	.02	.05	-.05	-.13	.06	.07	.09	-.05
RC2C3	**.60**	-.08	-.22	.07	**.44**	.11	.16	-.08	.00	-.14	-.02	.11	-.07	-.11	.05	.07	.06	-.04
RC2C4	**.52**	-.06	-.22	.12	**.41**	.08	.21	-.16	.02	-.08	-.03	.05	-.07	.04	-.04	.02	-.07	-.04

Item code	Factor 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
RC2C5	**.62**	-.11	-.29	.06	**.50**	.16	.14	-.12	.06	-.11	-.02	.06	-.08	-.06	.01	.00	.02	-.02
RC2C7	**.57**	-.14	-.20	.10	**.40**	.06	.09	-.17	.06	-.01	-.03	-.03	-.04	.02	-.07	.01	-.11	-.01
RD1A2	.29	.25	-.03	-.07	.00	-.06	.08	-.01	.12	-.17	-.07	.13	**.40**	.00	-.15	-.15	.10	.01
RD1A3	**.43** [a]	.15	-.05	-.10	-.03	-.14	.10	-.09	.12	-.05	-.11	.09	.25	.08	-.08	-.01	-.01	-.01
RD1A5	**.35** [a]	.22	.07	.03	-.02	-.01	.01	.02	.18	-.18	-.08	.22	.29	.03	-.10	-.15	.09	-.04
RD1A10	**.44** [a]	.19	.09	-.13	.02	-.05	.07	-.06	.09	-.07	-.07	.14	.12	-.05	-.11	-.04	.01	-.04
RD2A1	**.48** [a]	-.22	-.13	-.20	-.12	.20	.03	-.13	.06	.11	-.01	.02	.07	.02	.02	.00	-.10	.02
RD2A2	**.55** [a]	-.15	-.23	-.14	-.15	.27	.07	-.06	.03	.04	-.05	.02	-.05	.01	-.04	.00	.05	.00
RD2A5	**.55**	**-.31**	-.19	-.25	-.25	**.37**	-.01	-.02	.10	.05	.01	.04	.02	.02	.02	-.05	-.01	-.04
RD2A6	**.58**	-.28	-.22	-.25	**-.32**	**.39**	.02	-.02	.14	.02	-.07	.07	-.03	-.04	.01	.04	-.02	.01
RD2A8	**.56**	-.16	-.12	**-.30**	-.15	.27	-.02	.06	.09	.04	-.01	.02	-.05	-.01	.02	.03	.04	.05
RD2A10	**.58**	-.26	-.28	-.28	-.31	**.38**	.02	-.01	.12	.04	-.07	.07	.00	-.06	.01	.07	-.04	.03
RD3A1	**.59**	.18	-.13	.12	-.03	-.06	.23	-.07	.00	**.31**	.01	-.22	.11	-.08	-.04	-.03	.00	.12
RD3A3	**.56**	.22	-.13	.12	-.03	-.13	.26	-.12	.04	**.35**	-.04	-.19	.11	-.09	-.10	.02	-.05	.10
RD3A4	**.54** [a]	.27	.00	.03	-.16	-.22	.08	-.04	.10	-.03	.02	-.03	-.09	-.11	.09	.10	.12	-.02
RD3A7	**.44** [a]	.19	.00	-.07	-.04	-.03	.11	.06	.14	-.01	.03	-.16	-.01	-.05	.08	-.12	.18	.03
RD3A10	**.49** [a]	.14	-.26	.08	-.09	.00	.16	.00	.06	.08	.10	-.16	-.04	.01	-.10	-.18	.16	.00
RD3A12	**.57** [a]	.25	-.09	.11	-.13	-.16	.20	-.06	.06	.13	-.02	-.16	-.04	-.14	.03	.13	.19	-.11
RD3A14	**.64** [a]	.23	-.13	.07	-.10	-.12	.15	-.09	-.04	.14	-.03	-.08	.01	-.04	.06	.10	.01	-.09
RD4A2	**.38**	.26	**-.32**	-.02	-.09	.01	.01	.18	-.15	-.10	-.07	-.01	.02	-.07	-.07	-.02	-.05	-.07
RD4A3	**.54** [a]	.18	-.29	-.02	-.08	-.06	.07	.16	-.26	-.02	-.04	-.07	.03	.04	-.06	-.11	-.11	.02

	Factor																	
Item code	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
RD4A4	**.49** [a]	.11	-.19	-.02	-.10	-.09	.06	.09	-.07	-.10	.12	.07	-.07	-.03	.06	-.05	-.03	-.04
RD4A6	**.64** [a]	.22	-.22	.07	-.15	-.04	.10	.12	-.16	-.12	.04	-.01	-.16	-.01	-.02	-.17	-.13	-.01
RD4A7	**.68** [a]	.27	-.17	.01	-.17	-.09	.13	.09	-.09	-.14	.05	.01	-.12	-.05	.01	-.12	-.18	-.05

Note. Items with loadings above .30 in the primary factor and without cross-loadings on secondary factors were selected.

Loadings higher than .30 are marked in bold.

[i] Inverted item.

[a] Selected item.

Table 8

Selected and Discarded Items in Each Step of the Analysis

	Selected Items	**Discarded Items**
Step 1: Exploratory Item Analysis According Unidimensionality (ES1 sample)		
EFA	SA1A1, SA1A2, SA1A3, SA1A7, SA1B2, SA1B10[i], SA2A2, SA2A5, SA3A2i, SA3A3, SA3A7, SA3A8, SA3A9, SA3A11, SA3A13, SA3A14, SA3A16[i], SB1A2, SB1C6, SB2A4, SB4C2, SB4C7, SC1A2, SC1A3, SC1B3, SC1B5, SC2B1, SC3A2[i], SC3A4, SC3B5, SC3B6, SC4A5, SC4C1[i], SC4D1, SC4D7, RA1A2, RA1A3, RA1A5, RA1A6, RA1A7, RA1A8, RA2A2, RA2A10, RA3A3, RA3A4, RA3A6, RA4A5, RA4B2, RA4B3, RA4B6, RB1A8, RB1A9, RB2A1, RB2A2, RB3A3, RB3A5, RB3A8, RB3B2, RB3B3, RB3B12, RB3B13, RB4A2, RB4A6, RB5A2, RB5A9, RC2A4, RC2B5, RD1A3, RD1A5, RD1A10, RD2A1, RD2A2, RD3A4, RD3A7, RD3A10, RD3A12, RD3A14, RD4A3, RD4A4, RD4A6, RD4A7	SA1B3, SA1B4, SA2A3, SA2A17, SA2A1O[i], SAB1A7[i], SB1B2, SB1B4[i], SB1C7, SB1C9, SB2A1, SB2A2, SB2A6, SB2B1[i], SB2B2[i], SB2B5, SB3A6, SB3B1, SB3B2, SB3B3, SB4A1, SB4A9, SB4B1, SB4C3, SB4C6, SC1B1, SC1B4, SC3B1, SC3B2, SC3B7, SC4A4, SC4B3, SC4B4, SC410, SC4B12, SC4D5, RA2A3, RA2A5, RA2A8, RA3A1, RA3A2, RA3A8, RA4B5, RB1A5, RB1A11[i], RB1A12[i], RB2A6, RB3B4, RB2B6, RB2B7, RB4A1, RB4A13, RB4A15, RB5A3, RB5A4, RB5B8, RC1A2, RC1A4, RC1A5, RC1A7, RC1A9 RC1A11, RC2A3, RC2A6, RC2C2, RC2C3, RC2C4, RC2C5, RC2C7, RD1A2, RD2A5, RD2A6, RD2A8, RD2A10, RD3A1, RD3A3, RD4A2
Step 2. Multiple Group Confirmatory Factor Analysis According to Simple Factor Structure and Language Invariance		
Step2 a. EFA in the CFA framework (ES1 sample) Model1$_{\text{Fitted}}$ (Simple Factor Structure)	SA1A7, SA2A2, SA3A2[i], SA3A3, SA3A7, SA3A8, SA3A11, SA3A16[i], SB1C6, SB4C2, SC1A3, SC3A2[i], SC3B5, SC4A5, SC4C1[i], SC4D1, RA1A2, RA1A3, RA1A5, RA1A8, RA2A10, RA3A3, RA3A6, RA4A5, RA4B2, RB2A2, RB3B2, RB4A2, RC2B5, RD1A3, RD1A5, RD1A10, RD3A4, RD3A7, RD3A14, RD4A3, RD4A4	SA1A1, SA1A2, SA1A3, SA1B2, SA1B10[i], SA2A5, SA3A9, SA3A13, SA3A14, SB1A2, SB2A4, SB4C7, SC1A2, SC1B3, SC1B5, SC2B1, SC3A4, SC3B6, SC4D7, RA1A6, RA1A7, RA2A2, RA3A4, RA4B3, RA4B6, RB1A8, RB1A9, RB2A1, RB3A3, RB3A5, RB3A8, RB3B3, RB3B12, RB3B13, RB4A6, RB5A2, RB5A9, RC2A4, RD2A1, RD2A2, RD3A10, RD3A12, RD4A6, RD4A7
Step2b. Multiple Group Simple Factor Structure and Language Invariance (ES1 and EN samples)		

	Selected Items	**Discarded Items**
Step.2.b.Model2 (Model Constrained to Configural Invariance)	SA1A7, SA2A2, SA3A3, SA3A7, SA3A8, SA3A11, SA3A16[i], SC3B5, SC4C1[i], SC4D1, RA1A2, RA2A10, RA3A3, RA3A6, RB3B2, RC2B5, RD1A5, RD3A7, RD3A14	SA3A2[i], SB1C6, SB4C2, SC1A3, SC3A2[i], SC4A5, RA1A3, RA1A5, RA1A8, RA4A5, RA4B2, RB2A2, RB4A2, RD1A3, RD1A10, RD3A4, RD4A3, RD4A4
Step.2.b.Model3$_{MINV}$ (Model Constrained to Metric Invariance)	SA1A7, SA2A2, SA3A3, SA3A7, SA3A8, SA3A11, SC3B5, SC4C1[i], SC4D1, RA2A10, RA3A3, RA3A6, RB3B2, RC2B5, RD1A5, RD3A7, RD3A14	SA3A16[i], RA1A2
Step.2.b.Model3$_{PMINV}$ (Model resulted in Partial Matric Invariance)	SA1A7, SA2A2, SA3A3, SA3A7, SA3A8, SA3A11, SC3B5, SC4C1[i], SC4D1, RA2A10, RA3A3, RA3A6, RB3B2, RC2B5[a], RD1A5, RD3A7, RD3A14	
Step 2c. Content Reconstruction, Selection of Additional Items to Fill Content Gaps (ES1, EN, and ES2 samples)		
Network analysis	**Candidate Items Submitted to Network Analysis to be Reinserted in the Pool**	**Selected Items to be Reinserted in the Pool**
	SB2B2[i], SB2B1[i], SB4B1, SB4C2, RB1A5, RB2B4, RB4A13, RB4A2, RB5A3, RC2A3, RC2A6, RC2C3, RC2C5 Alternative pool of items SOCIAL-BAP: SRS12[i], SRS60; RIRE-BAP: AQ39[i], BAPQ17, SRS31	SB2B1[i], SB4B1, SRS60, RB4A2, RB4a13, RB5A3, SRS31, RC2A3, RC2C3
Multiple Group Simple Factor Structure and Language Invariance (ES1, EN, and ES2 samples)		
Step.2.c.Model2 (Simple Factor Structure, Model Constrained to Configural Invariance)	SA1A7, SA2A2, SA3A3, SA3A7, SA3A8, SA3A11, SB4B1[a], SRS60, SC3B5, SC4C1[i], SC4D1, RA2A10, RA3A3, RA3A6, RB3B2, RB4A2, RC2A3, RC2B5, RC2C3, RD1A5, RD3A7, RD3A14	SB2B1[i], RB4A13, RB5A3, SRS31
Step.2.b.Model2$_{MINV}$ (Model Constrained to Metric Invariance)	SA1A7, SA2A2, SA3A7, SA3A8, SA3A11, SB4B1, SRS60, SC3B5, SC4D1, RA2A10, RA3A3, RA3A6,	SA3A3, SC4C1[i]

	Selected Items	Discarded Items
	RB3B2, RB4A2, RC2A3, RC2B5, RC2C3, RD1A5, RD3A7, RD3A14	
Step.2.C.Model3$_{PMINV}$ (Model resulted in Partial Matric Invariance)	SA1A7, SA2A2, SA3A7, SA3A8, SA3A11, SB4B1[a], SRS60, SC3B5, SC4D1, RA2A10, RA3A3, RA3A6, RB3B2, RB4A2, RC2A3[a], RC2B5[a], RC2C3[a], RD1A5, RD3A7, RD3A14	
Step.2.c.Model4 (Model Constrained to Structural Invariance)	SA1A7, SA2A2, SA3A7, SA3A8, SA3A11, SB4B1[a], SRS60, SC3B5, SC4D1, RA2A10, RA3A3, RA3A6, RB3B2, RB4A2, RC2A3[a], RC2B5[a], RC2C3[a], RD1A5, RD3A7, RD3A14	

Note. This table shows specific items eliminated and selected on each stage and model as well as items reincorporated to the pool.

[i] Inverted item.

[a] Item behaving differently across groups (metric partial invariance).

Table 9
Factor Standardized Loadings from Step2.a.Model1$_{Fitted}$ Simple Factor Structure Item Selection (ES1 Sample)

Item code	Estimate loading in SOCIAL-BAP	Estimate loading in RIRE-BAP
SOCIAL-BAP		
SA1A7	-.76	-.07
SA2A2	.64	-.04
SA3A2[i]	-.56	.10
SA3A3	.60	.12
SA3A7	.64	-.08
SA3A8	.63	.06
SA3A11	.59	-.09
SA3A16[i]	-.61	.09
SB1C6	.47	.25
SB4C2	.34	.17
SC1A3	.48	.29
SC3A2[i]	-.83	.23
SC3B5[i]	.62	.06
SC4A5	.81	.00
SC4C1[i]	-.55	.02
SC4D1	.76	.04
RIRE-BAP		
RA1A2	-.00	.58
RA1A3	.09	.58
RA1A5	.11	.48
RA1A8	.15	.59
RA2A10	.07	.51
RA3A3	.06	.51
RA3A6	.01	.63
RA4A5	-.07	.48
RA4B2	-.14	.64
RB2A2	.16	.41
RB3B2	.08	.41
RB4A2	-.06	.41
RC2B5	.01	.48
RD1A3	.08	.42
RD1A5	-.06	.42
RD1A10	-.04	.45
RD3A4	-.02	.63
RD3A7	-.05	.50
RD3A14	.00	.69
RD4A3	.03	.53
RD4A4	.13	.45

Note. This table shows estimate loadings for selected items in *Step2.a.Model1$_{Fitted}$*.
[i] Inverted item.

Table 10

Adjacency Matrix for Content Reconstruction: Selection of Additional Items to Fill Content Gaps (ES1, EN, and ES2 samples)

	SB2b1i	SB2b2i	SB4b1	SB4c2	SRS12i	SRS60	RB1a5	RB2b4	RB4a2	RB4a13	RB5a3	RC2a3	RC2a6	RC2c3	RC2c5	AQ39i	BAPQ17	SRS31	SOCIAL-BAP
SB2b1i																			
SB2b2i	.60																		
SB4b1	.06	.03																	
SB4c2	_	.14	.01																
SRS12i	.15	.07	.07	.02															
SRS60	.03	.02	.18	.12	.29														
RB1a5	_	_	.05	-.11	-.07	_													
RB2b4	.03	_	.06	_	.07	.02	.17												
RB4a2	_	_	_	_	_	_	_	_											
RB4a13	_	_	.03	.13	_	_	-.05	_	.35										
RB5a3	.01	.03	.03	.16	_	_	.12	.14	_	.14									
RC2a3	_	_	_	_	_	.02	-.11	_	_	_	_								
RC2a6	_	_	_	_	_	.04	_	_	.03	.02	_	.31							
RC2c3	_	_	.01	_	_	.01	.04	.07	.09	-.09	_	.09	_						
RC2c5	-.03	-.04	_	_	_	_	.14	.02	_	_	_	_	.07	.51					
AQ39i	-.03	_	-.06	_	.06	-.05	_	_	_	-.05	_	_	_	_	_				
BAPQ17	.04	_	.09	.01	-.04	_	_	_	_	.07	.04	.01	.02	.04	_	-.60			
SRS31	-.09	-.04	.12	_	_	.02	.08	.12	.05	_	_	_	.03	.12	.08	-.09	.03		
SOCIAL-BAP	.11	_	.17	.09	.08	.21	.24	.18	-.13	_	_	_	.03	_	.01	_	_	.07	
RIRE-BAP	_	_	.09	.07	_	.05	.03	.03	.23	.06	.15	.19	.05	.09	.09	-.07	.05	.14	.21

Note. Partial correlation values of the items represented in the initial empirical network. Items were selected according to their relation with one of the BAP-IT dimensions. Further selection criteria can be consulted in the main manuscript.

Underscores represent the absence of relation between two items.

Figure 1a.

Initial Empirical Network Model for Content Reconstruction of Step.2.c

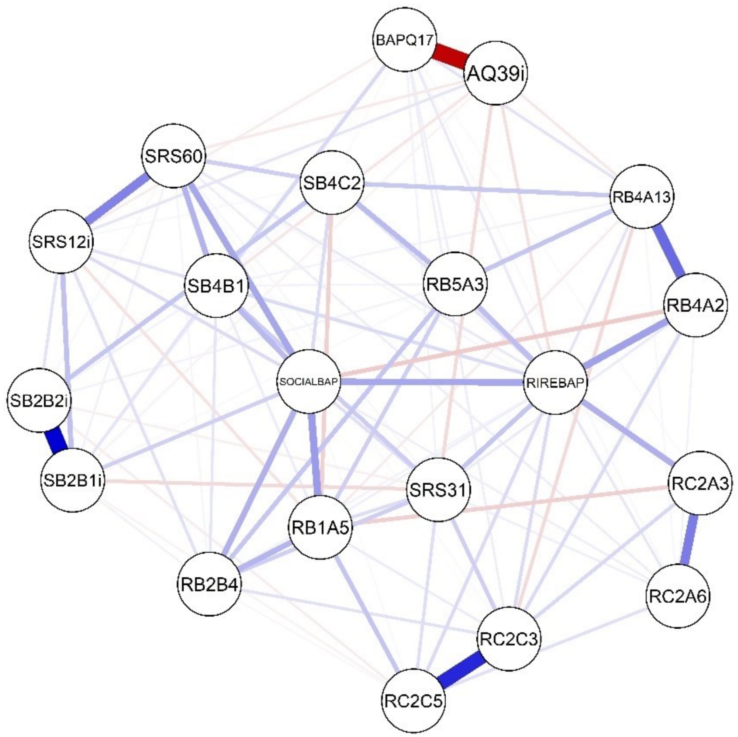
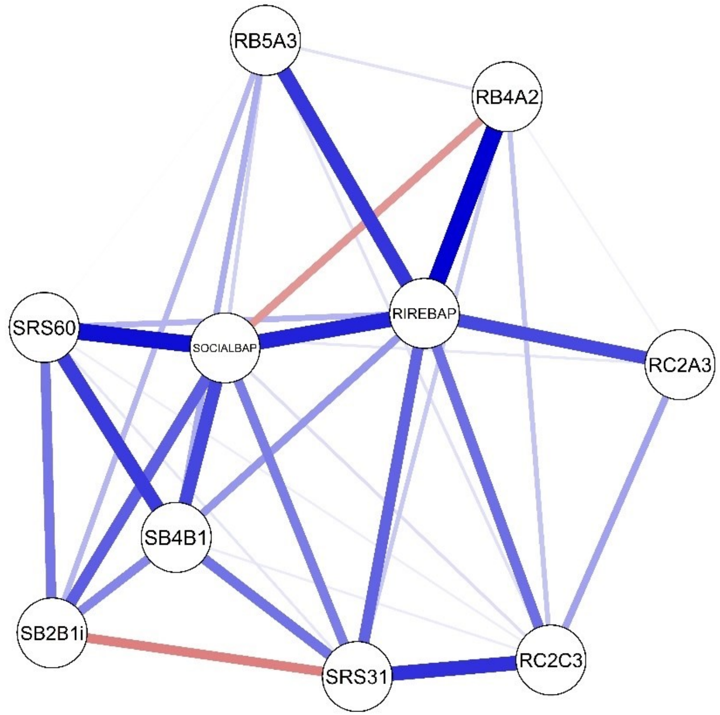


Figure 1b.

Final Empirical Network Model for Content Reconstruction of Step.2.c



Appendix 5. Chapter 6. Study 4.

"The Broad Autism Phenotype - International Test (BAP-IT): a Two-Domain Based Test for the Assessment of the Broad Autism Phenotype"

Document B. EFA in the CFA Framework (E/CFA) to Select Items Which Took Part in Alternative Pool of Items

Data Analysis

An EFA in the CFA framework (Brown, 2015; Jöreskog & Sörbom, 1979) was conducted with *maximum likelihood estimation with robust standard errors* (MLR) as the model estimator (Muthén & Asparouhov, 2011). The purpose of this analysis was to select the items which would form part of the *Alternative Pool of Items* in our study. Factor variances were fixed to one, factor means were fixed to zero, and factor covariance was freely estimated. Cross-loadings were fixed to zero for anchor items (AQ44 was the anchor item for SOCIAL-BAP and AQ39 for RIRE-BAP) and freed for the rest of the items. Fit indices and acceptable parameters were established at: RMSEA (≤ 0.06), CFI (≥ 0.95), and TLI (≥ 0.95) following Hu and Bentler (1999). The model comprehended the resulting 48 items of Godoy-Giménez et al. (2021a) work. Selected ten items were those with loadings above .40 in their respective dimension and with very low or null loadings on the other dimension (below 0.20; selected and discarded items can be consulted in Table 11). The dataset contained 349 cases from a single Spanish group. There were 23 missing values, full information maximum likelihood was used to estimate missing values (Muthén & Asparouhov, 2011). Analyses were performed in MPlus v7.11. (Muthén & Muthén, 2012).

Table 11

Item Selection in E/CFA Analysis of the Candidate Pool of Items (Godoy-Giménez et al., 2021a[19])

Dimension experts assigned the item to	**Items code**	**Estimate loading in SOCIAL-BAP**	**Estimate loading in RIRE-BAP**
SOCIAL-BAP	AQ1	.37	-.07
SOCIAL-BAP	AQ10	.29	.04
SOCIAL-BAP	AQ20	-.06	-.10
SOCIAL-BAP	AQ27	.12	.11
SOCIAL-BAP	AQ31	.14	.02
SOCIAL-BAP	AQ36	.14	.05
SOCIAL-BAP	AQ40	.19	.07
SOCIAL-BAP	**AQ44**	**.61** [a]	.00
SOCIAL-BAP	AQ45	-.17	-.07
SOCIAL-BAP	**AQ47**	**.54** [a]	-.02
SOCIAL-BAP	AQ50	.12	-.02
SOCIAL-BAP	**BAPQ1**	**-.70** [a]	.09
SOCIAL-BAP	BAPQ10	.09	.05
SOCIAL-BAP	BAPQ21	.01	-.06
SOCIAL-BAP	BAPQ34	-.19	-.13
SOCIAL-BAP	**BAPQ36**	**-.86** [a]	.06
SOCIAL-BAP	SRS10	.11	.39
SOCIAL-BAP	**SRS12**	**.42** [a]	-.05
SOCIAL-BAP	SRS13	.28	.15
SOCIAL-BAP	SRS2	.29	.18
SOCIAL-BAP	SRS26	-.20	.01
SOCIAL-BAP	SRS35	.24	.11
SOCIAL-BAP	SRS37	.38	.14
SOCIAL-BAP	SRS38	-.11	-.18
SOCIAL-BAP	SRS47	.10	.33
SOCIAL-BAP	SRS5	-.04	.19
SOCIAL-BAP	SRS53	.15	.08
SOCIAL-BAP	**SRS60**	**.50** [a]	.02
SOCIAL-BAP	**SRS64**	**.63** [a]	.03
SOCIAL-BAP	SRS7	-.18	-.10
RIRE-BAP	AQ12	.08	-.13
RIRE-BAP	AQ2	-.14	-.22
RIRE-BAP	AQ25	.16	.31
RIRE-BAP	**AQ39**	.00	**-.47** [a]
RIRE-BAP	AQ41	-.04	-.20
RIRE-BAP	AQ46	-.41	-.27

[19] Those results represent ulterior analysis that were not included in the published paper.

Dimension experts assigned the item to	Items code	Estimate loading in SOCIAL-BAP	Estimate loading in RIRE-BAP
RIRE-BAP	BAPQ15	-.20	-.37
RIRE-BAP	**BAPQ17**	-.26	**.56** [a]
RIRE-BAP	BAPQ22	.32	.59
RIRE-BAP	BAPQ24	.20	.52
RIRE-BAP	BAPQ33	.08	.29
RIRE-BAP	SRS20	.09	.31
RIRE-BAP	SRS24	.28	.34
RIRE-BAP	SRS28	.06	.39
RIRE-BAP	**SRS31**	.12	**.47** [a]
RIRE-BAP	SRS42	.07	.35
RIRE-BAP	SRS50	.20	.25
RIRE-BAP	SRS61	.17	.24

Note. Table contains all the items chosen by the experts in Godoy-Giménez et al. (2021a) manuscript.
Item codes represented the original questionnaire of the items.
Loadings of the selected items in the respective factor is displayed in bold.
[a] Selected item.