



ADAPTACION DE DOS INSTRUMENTOS PARA NIÑOS COLOMBIANOS: LA ESCALA BARRATT DE IMPULSIVIDAD (BIS-11C) Y EL CUESTIONARIO DE AGRESIVIDAD DE BUSS Y PERRI (AQ)

Nicolás Chahín Pinzón

Dipòsit Legal: T.1557-2013

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.

Nicolás Chahín Pinzón

**ADAPTACIÓN DE DOS INSTRUMENTOS PARA NIÑOS
COLOMBIANOS: LA ESCALA BARRATT DE
IMPULSIVIDAD (BIS-11c) Y EL CUESTIONARIO DE
AGRESIVIDAD DE BUSS Y PERRY (AQ)**

TESIS DOCTORAL

Dirigida por

Dr. Andreu Vigil Colet

Dr. Urbano Lorenzo Seva

Departamento de Psicología



UNIVERSITAT ROVIRA I VIRGILI

Tarragona

2013

Dedicatoria

A mi esposa Blanca

A mi madre Carmen

A mi padre Atala

A mi hijo Juan Nicolás

A mi hijo Julián Eduardo

Agradecimientos

Al Dr. Andreu Vigil Colet

Al Dr. Urbano Lorenzo Seva

Por toda su colaboración y apoyo

ÍNDICE

INTRODUCCIÓN	14
1 La Escala Barratt de Impulsividad	15
1.1 Conceptualización de impulsividad	15
1.2 Evolución y desarrollo de la escala	17
1.3 La versión del test BIS-11	19
1.4 Transculturalidad del test BIS-11	23
1.5 La estructura factorial del test BIS-11 en adultos	24
1.6 Adaptaciones del test BIS-11 en niños y adolescentes	30
1.6.1 La adaptación italiana para adolescentes: BIS-11A	34
1.6.2 La adaptación española para niños: BIS-11c	36
2 El Cuestionario de Agresividad de Buss y Perry	40
2.1 Conceptualización de agresividad	40
2.2 Los orígenes: The Hostility Inventory de Buss y Durkee	41
2.3 Desarrollo del test AQ	44
2.4 Transculturalidad del test AQ	46
2.5 El modelo del test AQ en adultos	47
2.5.1 Estudios que informan del modelo tetrafactorial	47
2.5.2 Estudios que informan diferentes factores	52
2.5.3 Las versiones refinadas del test AQ	53
2.5.4 La versión reducida española del test AQ consistente entre culturas	57
2.6 Adaptaciones del test AQ en niños y adolescentes	61
3 La adaptación de test entre culturas	66
3.1 Equivalencia	67
3.2 Sesgos en el proceso de adaptación	67

3.2.1 Sesgos de constructo	68
3.2.2 Sesgos de método	69
3.2.3 Sesgos de los ítems.....	69
3.3 Directrices para la adaptación de test	71
OBJETIVOS	73
1 Objetivos relacionados con el test BIS-11c	74
2 Objetivos relacionados con test AQ	75
MÉTODO	78
1 Participantes	79
2 Instrumentos	81
2.1 El test BIS-11c	81
2.2 El test AQ	81
3 Procedimiento para la adaptación de los instrumentos	82
3.1 Procedimiento para la adaptación del test BIS-11c	82
3.2 Procedimiento para la adaptación del test AQ	84
4 Procedimiento para la administración de los dos instrumentos	86
5 Procedimiento para el análisis de datos.....	87
5.1 Procedimiento para el análisis de datos del test BIS-11c	87
5.2 Procedimiento para el análisis de datos del test AQ	89
RESULTADOS	92
1 Resultados referentes al test BIS-11c	93
2 Resultados referentes al test AQ	103
DISCUSIÓN	116
1 Discusión de la adaptación del test BIS-11c.....	117

1.1 Sobre la estructura factorial del test BIS-11c	117
1.2 Aspectos concernientes a las comparaciones multigrupo	122
1.3 Consistencia interna del test BIS-11c	126
1.3.1 Las escalas impulsividad motora e impulsividad no planificadora	126
1.3.2 La escala impulsividad cognitiva.....	128
2 Discusión de la adaptación del test AQ.....	130
2.1 Sobre la estructura factorial del test AQ	130
2.2 Consistencia interna del test AQ	134
2.2.1 Comparación de la consistencia interna con las versiones para adultos	134
2.2.2 Comparación de la consistencia interna con las versiones para niños y adolescentes.....	138
2.2.3 Consideraciones finales concernientes a la consistencia interna	143
2.3 Consideraciones sobre las diferencias de agresividad entre sexos	145
CONCLUSIONES	148
1 Conclusiones para la adaptación del test BIS-11c.....	149
2 Conclusiones para la adaptación del test AQ.....	150
REFERENCIAS BIBIOGRÁFICAS	152

INTRODUCCIÓN

1 La Escala Barratt de Impulsividad

1.1 Conceptualización de impulsividad

En las últimas décadas, hay un interés cada vez mayor por el estudio de la impulsividad ya que se encuentra relacionada con diversos tipos de trastornos, tanto en la niñez como en la adultez. Si bien existe acuerdo sobre su relevancia clínica, su definición ha sido caracterizada por una falta de consenso (Eysenck y Eysenck, 1977; Arce y Santisteban, 2006). En lenguaje común la palabra *impulso* es definida como un deseo o motivo de hacer algo súbitamente y sin premeditación. La Real Academia Española define la palabra *impulsivo* como aquella persona que habla o actúa sin reflexión y/o cautela, dejándose llevar por la impresión del momento (Real Academia Española, 2001). En general se puede decir que la impulsividad es un constructo amplio que incluye aspectos cognitivos y conductuales, compuesto por diferentes factores (Evsenden, 1999).

La impulsividad ocupa un papel significativo en las teorías de la personalidad (Eysenck y Eysenck, 1985; Cloninger, Przybeck y Svrakic, 1991). Se le ha intentado definir de distintas formas, como por ejemplo, la dificultad para sostener la atención o la tendencia a actuar dejándose influenciar por el ímpetu del momento, sin tener presente los potenciales riesgos (Eysenck, Pearson, Easting y Allsopp, 1985), la predisposición a responder, más que a inhibir la respuesta (Buss y Plomin, 1975) o la

propensión a actuar con una menor reflexión en comparación con los individuos del mismo nivel de habilidad y conocimiento (Dickman, 1990).

Barratt (1993) ha explicado la impulsividad teniendo en cuenta cuatro elementos: biológico, cognitivo, ambiental y comportamental. La impulsividad para él es una dimensión de primer orden de la personalidad que se encuentra esencialmente relacionada con el control de impulsos. Propone una definición clínica, expresada como una “predisposición” para reaccionar de forma rápida y no planeada ante estímulos internos o externos, sin tener en cuenta las consecuencias negativas que las conductas impulsivas tengan para el individuo o para los otros. La definición anterior tiene varias implicaciones. Primera, la impulsividad es vista como una predisposición, y en consecuencia es considerada como parte de un patrón conductual. Segunda, la impulsividad involucra una acción rápida y no planeada. Tercera, la impulsividad está asociada con actuar sin tener presente las consecuencias que acarrearán estas acciones.

Para los efectos de este estudio, se considerará la impulsividad como una predisposición a reaccionar de manera rápida y no planeada a estímulos, tanto internos como externos, sin tener en cuenta las consecuencias negativas, para sí mismo como para los demás (Barratt, 1985, 1993; Barratt y Patton, 1983; Eysenck y Eysenck, 1978; Moeller, Barratt, Dougherty, Schmitz y Swann, 2001).

1.2 Evolución y desarrollo de la escala

En los inicios de su carrera, a partir de la teoría del aprendizaje, E. S. Barratt estudió la relación entre ansiedad (fuerza del hábito) e impulsividad (oscilación conductual). En sus experimentos aplicó a estudiantes universitarios escalas de ansiedad e impulsividad provenientes del *Thurstone Temperament Schedule* y *Guilford-Zimmerman Temperament Survey*, y estos datos los comparó con el desempeño en tareas cognitivas y motoras. Los resultados confirmaron que tanto ansiedad como impulsividad correlacionaban con el desempeño de las actividades realizadas. En especial, un pobre resultado estaba frecuentemente asociado con la complejidad de la tarea. Y los sujetos con alta impulsividad exhibían problemas en planeación, ejecución y desempeño perceptual motor (Barratt, 1994). Pero también sus hallazgos demostraron que impulsividad y ansiedad tenían muy bajas correlaciones entre sí, comprobando que en ellas estaban implicados diferentes sistemas neuronales, y además, presentaban distintos correlatos biológicos y conductuales con otras variables de la personalidad. Estos resultados lo llevaron a teorizar que las dimensiones impulsividad y ansiedad eran ortogonales (Barratt, 1965, 1994; Patton y Stanford, 2011).

El desarrollo de la *Barratt Impulsiveness Scale* (en adelante BIS) tiene su inicio a finales de los años cincuenta y ha sido objeto de un largo proceso de transformación. De hecho hay once revisiones a través de su historia, a partir del año 1959 hasta la última versión de Patton, Stanford y Barratt (1995), en donde la concepción de las

dimensiones ha ido variando, y por consiguiente, el nombre y número de los factores, la cantidad de ítems y el tipo de respuesta. En un principio se utilizaron ítems originales de otras escalas, luego, el proceso se fue depurando con los años, en especial porque se quería obtener un instrumento que presentara bajas correlaciones con la ansiedad. Ya en el año 1965, con un test más elaborado, el BIS-5, se encontraron cuatro factores: *velocidad en la respuesta cognitiva*, *pérdida de control del impulso*, *búsqueda de aventura o extroversión* y *toma de riesgos* (Barratt, 1965, 1994).

En los años sucesivos el BIS tuvo otras revisiones, y en 1983 después de años de investigación, se hipotetizó que la impulsividad estaba compuesta por tres factores: *impulsividad motora* que implicaba actuar sin pensar, *impulsividad cognitiva* que se relacionaba con la toma de decisiones rápidas, e *impulsividad no planificadora*, caracterizada por una orientación hacia el presente o una falta de planificación hacia el futuro. Estas tres escalas se encontraban en la nueva versión del BIS-10, que poseía 34 ítems. De *impulsividad motora* 11 ítems, de *impulsividad cognitiva* 11 ítems y de *impulsividad no planificadora* 12 ítems. Los coeficientes de fiabilidad estuvieron entre $\alpha=.89$ y $\alpha=.92$ (Barratt, 1985,1994).

Pero la estructura factorial del BIS-10 presentaba dificultades de replicación, y ante estos problemas se realizaron varias investigaciones no publicadas. En una de ellas el análisis factorial arrojó tres factores, el primero, *ideo-motor*, que contenía ítems de la anterior versión de las escalas *impulsividad cognitiva* e *impulsividad motora*, el segundo, *planificación cuidadosa*, y el tercero, *estabilidad en el*

afrentamiento, que combinaba las escalas *orientación futura* y *estabilidad de afrontamiento en la vida diaria* (Barratt, 1994; Parker y Bagby, 1997).

Por su parte Luengo, Carrillo de la Peña y Otero (1991) en una adaptación española del BIS-10 lo aplicaron a una muestra de estudiantes universitarios y no pudieron identificar la estructura de tres factores, y además, hallaron coeficientes de fiabilidad bastante bajos en todas las escalas, en especial para *impulsividad cognitiva*. Según expone Barratt (1994), esta dificultad es debida a que las funciones cognitivas son difíciles de medir por su carácter inferencial. Sin embargo, otro estudio que utilizó el BIS-10, sí pudo identificar el factor *impulsividad cognitiva*, dentro de una solución de 15 factores (Gerbing, Ahadi y Patton, 1987).

1.3 La versión del test BIS-11

Posteriormente Patton *et al.* (1995) en un intento por mejorar el instrumento, compararon el BIS-10 con una nueva versión, en tres muestras: estudiantes universitarios ($N=412$), pacientes psiquiátricos ($N=248$) e internos penitenciarios ($N=164$). Su propósito fue revisar la estructura factorial del BIS-10 y comparar los resultados con la versión designada como BIS-11. Para lo cual se tuvo como norma eliminar del análisis los ítems que tuvieran una baja correlación ítem total, o que no contribuyeran a la diferenciación entre los rasgos extremos de los grupos. Se retiraron entonces los ítems 19 (*I have regular health check ups*), 26 (*I walk and move fast*), 27 (*I solve problems by trial-and-error*) y 29 (*I talk fast*). Con los restantes 30 ítems se llevó a cabo un análisis factorial exploratorio de componentes principales, y se

identificaron seis factores de primer orden en la rotación Promax. Éstos fueron denominados como se enuncia a continuación. El primero, *atención*, definido como centrarse en la tarea o actividad que se está realizando, compuesto por los ítems: *I don't pay attention*, *I concentrate easily**, *I "squirm" at plays or lectures*, *I am a steady thinker** y *I am restless at the theater or lectures* (el asterisco indica que es invertido). El Segundo, *impulsividad motora*, definido como actuar pensando en el afán del momento, compuesto por los ítems: *I do things without thinking*, *I make-up my mind quickly*, *I am happy-go-lucky*, *I act on impulse*, *I act on the spur of the moment*, *I buy things on impulse* y *I spend or charge more than I earn*. El tercero, *autocontrol*, definido como planear y pensar con detenimiento, compuesto por los ítems: *I plan tasks carefully**, *I plan trips well ahead of time**, *I am self controlled**, *I am a careful thinker**, *I plan for job security** y *I say things without thinking*. El cuarto, *complejidad cognitiva*, definido como disfrute de actividades que impliquen reto mental, compuesto por los ítems: *I save regularly**, *I like to think about complex problems**, *I get easily bored when solving thought problems*, *I am more interested in the present than the future* y *I like puzzles**. El quinto, *perseverancia*, definido como un estilo de vida consistente, compuesto por los ítems: *I change jobs*, *I change residences*, *I can only think about one thing at a time* y *I am future oriented**. Y por último, el sexto, *inestabilidad cognitiva*, definido como la tendencia a la distracción debida a la presencia de pensamientos veloces o intrusivos, compuesto por los ítems: *I have "racing" thoughts*, *I change hobbies* y *I often have extraneous thoughts when thinking*.

Utilizando rotaciones oblicuas, los factores de primer orden dieron lugar a tres de segundo orden. El primero, *impulsividad motora* con 11 ítems, incluía *impulsividad motora* y *perseverancia*. El segundo, *impulsividad no planificadora* con 11 ítems, incluía *autocontrol* y *complejidad cognitiva* (Patton *et al.*, 1995). Es conveniente señalar que estos dos factores estuvieron acordes a lo reportado en la investigación previa de Luengo *et al.* (1991) y eran consistentes con los originalmente hipotetizados para el BIS-10, y al mismo tiempo se encontraban relacionados con las dimensiones de impulsividad propuestas por Eysenck y Eysenck (1977). El tercero, denominado *impulsividad atencional* con 8 ítems, comprendía los factores *atención* e *inestabilidad cognitiva*. Este último no fue consistente con la posición teórica anterior de Barratt (1985) que postulaba el factor *impulsividad cognitiva*.

El hecho de que los ítems de este factor tuvieran cargas en todos los demás factores, implica para Barratt (1994) que el proceso cognitivo subyace a la impulsividad en general. Hay algunas posibles razones que intentan explicar por qué no pudo ser identificado el factor *impulsividad cognitiva*. La primera, que el aspecto cognitivo de la impulsividad no existe como tal, ya que el proceso de pensamiento subyace a todo el rasgo de impulsividad. La segunda, es posible que los sujetos en este instrumento no puedan medir independientemente el proceso de pensamiento implicado en la impulsividad (Barratt, 1994; Patton y Stanford, 2011; Patton *et al.*, 1995).

La correlación resultante entre los factores de primer orden estuvo entre $r=.15$ y $r=.42$ y en los de segundo orden entre $r=.46$ y $r=.53$. La correlación entre BIS-10 y

el BIS-11 fue de $r=.98$. El instrumento tuvo una buena consistencia interna total para las muestras de universitarios ($\alpha=.82$), drogadictos ($\alpha=.79$), pacientes psiquiátricos ($\alpha=.83$) e internos penitenciarios ($\alpha=.80$). No se encontraron diferencias entre las muestras estudiadas ($F(3, 657)=27.49$; $p<.0001$) y tampoco hubo diferencias de sexo ($F(2, 348)=17.58$; $p<.0001$) (Patton *et al.*, 1995).

Esta versión de Patton *et al.* (1995) ha sido empleada como escala de medida para evaluar muestras clínicas con diferentes tipos de psicopatologías. Una de las poblaciones a las que ha sido aplicado con mucha frecuencia el BIS-11, es aquella con trastornos por consumo de sustancias. Se ha utilizado en investigaciones como las de tabaquismo (Skinner, Aubin y Berlin, 2004), alcohol (Dom, D'haene, Hulstijn y Sabbe, 2006), cocaína (Moeller, Dougherty, Barratt, Schmitz, Swann y Grabowski, 2001) y éxtasis (Bond, Verheyden, Wingrove y Curran, 2004). El instrumento se ha administrado a diversos perfiles de consumidores, obteniendo éstos puntajes más altos en comparación con muestras normales. Igualmente los valores de las escalas son predictores de la cantidad de sustancia psicoactiva consumida (Forcada, Pardo y Bondia, 2006; Lane, Moeller, Steinberg, Buzby y Kosten, 2007; Lejuez, Bornovalova, Reynolds, Daughters y Curtin, 2007). Y en adicciones no farmacológicas como la ludopatía, los puntajes exhibidos son superiores a los grupos de control (Rodríguez-Jiménez *et al.*, 2006).

Del mismo modo el instrumento se ha administrado en otros estudios que comprenden enfermedades categorizadas en el Eje I del DSM-IV, como los relacionados con el estado de ánimo, tales como depresión y trastorno afectivo

bipolar. En general los puntajes correlacionan con los trastornos del estado de ánimo, adicionalmente, se reportan valores superiores en bipolares (Peluso *et al.*, 2007; Swann, Pazzaglia, Nicholls, Doughert y Moeller, 2003; Swann, Steinberg, Lijffijt y Moeller, 2008). En cuanto a los pacientes con riesgo de suicidio, se ha visto que presentan niveles mayores en *impulsividad motora* y *atencional*, en comparación con muestras sanas (Dougherty, Mathias, Marsh, Papageorgiou, Swann y Moeller, 2004). Otra de las enfermedades que exhibe altos puntajes en la escala, es el Trastorno por déficit de atención con hiperactividad (TDAH), uno de los que más asociados a la impulsividad y a los mecanismos de inhibición de la conducta, tanto en la niñez como en la adultez. Los resultados muestran que en estos pacientes hay déficit en los componentes motor, cognitivo y atencional (Malloy-Diniz, Fuentes, Leite, Correa y Bechara, 2007). Por otra parte, en el Eje II ha sido utilizado, entre otros, en la evaluación de trastornos de personalidad (Henry, Mitropoulou, New, Koenigsberg, Silverman y Siever 2001; Morgan, Gray y Snowden, 2011). Además, el BIS-11 ha sido recurrentemente aplicado a población carcelaria. Los individuos con historial criminal violento puntúan más alto en comparación con los grupos control, o con convictos detenidos por otro tipo de delitos (Folino, Escobar-Córdoba y Castillo, 2006; Smith, Waterman y Ward, 2006).

1.4 Transculturalidad del test BIS-11

Si bien las características de las muestras y los procedimientos psicométricos que se han seguido varían, el BIS-11 ha sido traducido y adaptado desde su versión

original inglesa a distintos idiomas y culturas, como ha sido reportado por Stanford, Mathias, Dougherty, Lake, Anderson y Patton (2009). Hay, entre otras, adaptaciones al alemán (Preuss *et al.*, 2008), chino (Yang, Yao y Zhu, 2007), coreano (Chung y Lee, 1997), francés (Baylé *et al.*, 2000), griego (Giotakos, Markianos, Vaidakis y Christodoulou, 2003), holandés (Goudriaan, Oosterlaan De Beurs y Van den Brink, 2008), italiano (Fossati, Di Ceglie, Acquarini y Barratt, 2001), japonés (Someya *et al.*, 2001), portugués (Von Diemen, Szobot, Kessler y Pechansky, 2007), turco (Güleç *et al.*, 2008) e indú (Singh, Solanki y Bhatnagar, 2008).

En idioma español igualmente existen diversas versiones como la española (Oquendo, Baca-García, Graver, Morales, Montalban y Mann, 2001), argentina (Folino *et al.*, 2006) y colombiana (Orozco-Cabal, Rodríguez, Herin, Gempeler y Uribe, 2010).

1.5 La estructura factorial del test BIS-11 en adultos

La identificación de la estructura factorial del BIS-11 ha sido problemática, de la misma manera que lo fue para el BIS-10, y la gran mayoría de los estudios no han podido replicar exactamente el modelo propuesto por Patton *et al.* (1995).

Lo que indican las diferentes investigaciones no es concluyente, y aunque existe evidencia que respalda una estructura multidimensional de la impulsividad, los componentes no son del todo consistentes, y es difícil encontrar una nosología común entre todas las muestras (Whiteside y Lynam, 2001). Para la verificación de su

estructura factorial en general han primado los estudios de carácter exploratorio sobre los confirmatorios. Empero, éstos últimos de hecho son más apropiados cuando se quiere verificar una estructura previa, brindando una mejor y más precisa información, por lo que es conveniente utilizar los modelos de ecuaciones estructurales para evaluar la bondad de ajuste del modelo (Byrne, 2008).

Puede afirmarse de manera general que la escala *impulsividad no planificadora* es la que tiende a ser encontrada más frecuentemente. Por otro lado, la escala más inestable es *impulsividad atencional*, ya que sus ítems cargan en diferentes dimensiones (Stanford *et al.*, 2009; Vasconcelos, Malloy-Diniz y Correa, 2012).

En primer término se encuentran aquellas investigaciones que intentan confirmar la estructura factorial propuesta de Patton *et al.* (1995). Entre ellas está la de Hou, Xiao, He, Li y Lui (2006) quienes adaptaron la versión inglesa al idioma chino en una muestra de 261 estudiantes universitarios. Eliminaron en el proceso, cuatro ítems de la escala original por su bajas cargas factoriales. En el análisis factorial los índices de ajuste al modelo fueron aceptables para *RMSEA*, más los otros indicadores estuvieron por debajo del punto de corte (*GFI*=.80, *AGFI*=.77, *RMSEA*=.069). La fiabilidad para la escala total fue de $\alpha=.75$, considerada como buena, pero en cambio, las escalas no tuvieron un buen comportamiento con valores para *impulsividad atencional* de $\alpha=.56$, *impulsividad motora* de $\alpha=.65$ e *impulsividad no planificadora* de $\alpha=.68$. De igual modo, Someya *et al.* (2001) con una muestra de 450 trabajadores y estudiantes, adaptaron sin eliminar ningún ítem el BIS-11 al idioma japonés. Del análisis confirmatorio informaron los siguientes dos indicadores:

$GFI=.85$ y $AGFI=.82$, que sugieren un ajuste pobre, pues se encuentran por debajo del límite de $.90$. La fiabilidad total fue de $\alpha=.80$, pero la consistencia de las escalas estuvo baja, *impulsividad atencional* con $\alpha=.60$, *impulsividad motora* con $\alpha=.64$ e *impulsividad no planificadora* con $\alpha=.65$. Y la de Preuss *et al.* (2008), que con sujetos sanos y psiquiátricos no pudieron confirmar la estructura original de Patton *et al.* (1995) para los seis factores de primer orden propuestos ($NFI=.61$, $RMSEA=.07$), obteniendo una consistencia interna para la escala total de $\alpha=.69$.

A nivel exploratorio tenemos entre muchas la adaptación italiana que utilizó una muestra de 700 estudiantes universitarios. Aunque se hallaron los tres factores, las cargas de los ítems fueron distintas a la versión inglesa. Un factor de segundo orden y cinco de primer orden fueron replicados adecuadamente. Aquí dos factores no correspondían al original, uno de ellos combinaba los factores de primer orden *atención* e *impulsividad motora*, y el otro, estaba conformado por *perseverancia* y *carencia en la postergación de la gratificación*. Y sólo el factor *impulsividad no planificadora* pudo ser replicado combinando *impulsividad cognitiva* y *autocontrol*. Estas discrepancias entre la versión italiana y la original fueron atribuidas más a aspectos culturales que a las diferencias entre las muestras, ya que ambas poseían similares niveles académicos y la distribución de sexos era similar. La consistencia interna total fue de $\alpha=.79$ (Fossati *et al.*, 2001). También la versión turca de Güleç *et al.* (2008) utilizando un análisis con rotación Promax en una muestra de 237 estudiantes universitarios y 83 pacientes psiquiátricos, encontraron tres factores, pero los ítems cargaron en factores distintos al original. La fiabilidad total fue buena tanto para los estudiantes como para los pacientes ($\alpha=.78$ y $\alpha=.81$), así como para la escala

impulsividad atencional ($\alpha=.76$ y $\alpha=.76$). Es de notar los bajos coeficientes en la escala *impulsividad motora* ($\alpha=.38$ y $\alpha=.27$) y en la escala *impulsividad no planificadora* ($\alpha=.23$ y $\alpha=.43$). Además de las señaladas, hay un sinnúmero de investigaciones a nivel exploratorio que informan sobre estructuras trifactoriales compuestas por ítems distribuidos en factores diferentes a los originales (Baylé *et al.*, 2000, Stanford *et al.*, 2009; Vasconcelos *et al.*, 2012).

Por otra parte, se encuentran las versiones refinadas que replican los tres factores con ligeras diferencias en la distribución de las cargas con respecto a la estructura original. Spinella (2007) desarrolló una versión de 15 ítems con una muestra de 700 adultos. Los 30 ítems de la versión original fueron sometidos a análisis de componentes principales, y la solución de tres factores fue especificada a priori. La gran mayoría de los ítems presentaron saturaciones dominantes en el factor esperado. Posteriormente se seleccionaron los cinco ítems de cada factor que presentaban las cargas más elevadas. Es necesario señalar que esta versión corta posee tres ítems que no se encuentran en las mismas escalas de la versión original. La consistencia interna para toda la muestra fue de $\alpha=.81$, no se reporta la fiabilidad para cada uno de los factores. Esta versión correlacionó significativamente con la de 30 ítems ($r=.94$). Y en esta dirección Orozco-Cabal *et al.* (2010) en Colombia, tomaron los mismos ítems seleccionados por Spinella (2007) con una muestra de 447 individuos, y administraron la versión española de Oquendo *et al.* (2001). La consistencia total informada fue de $\alpha=.79$. Asimismo, existe una adaptación de esta versión en alemán, que obtuvo valor de $\alpha=.81$ (Meule, Vögele y Kübler, 2011).

Y a nivel confirmatorio, aunque no se ha podido confirmar la estructura original, sí que se han propuesto otras soluciones de tres factores, en las que incluso se han asignado nombres nuevos a las dimensiones. Ireland y Archer (2008) con internos e internas penitenciarias, en el análisis confirmatorio no pudieron comprobar el modelo propuesto por Patton *et al.* (1995), de hecho los indicadores fueron insuficientes tanto para hombres ($RMSEA=.09$, $GFI=.77$, $CFI=.63$, $ECVI=3.82$), como para mujeres ($RMSEA=.09$, $GFI=.75$, $CFI=.62$, $ECVI=4.75$). Ante este resultado, utilizaron un análisis exploratorio de componentes principales con rotación Varimax, que produjo una solución con tres factores, distintos al modelo original. El primero, denominado *habilidades cognitivas y de planeación* ($\alpha=.85$), estaba compuesto principalmente por ítems de la escala *impulsividad atencional* e *impulsividad no planificadora*. Esto sugiere que el fracaso en la planeación incorpora la habilidad atencional. El segundo, *impulsivo motor comportamental* ($\alpha=.79$), formado por ítems mayormente de la escala *impulsividad motora* y algunos de la escala *impulsividad no planificadora*. Este factor posee alguna consistencia con la escala *impulsividad motora* de la versión original. Y el tercero, *distractibilidad* ($\alpha=.67$), con ítems de la escala *impulsividad motora* e *impulsividad atencional*. Este nuevo modelo de tres factores sólo tuvo un buen ajuste en hombres ($RMSEA=.047$, $GFI=.97$, $CFI=.98$, $ECVI=.32$), más en mujeres no funcionó. Los autores señalan de acuerdo a su evidencia, que al menos para este tipo de muestras carcelarias, el constructo de impulsividad no es estable entre los sexos. Resultado contrario a lo reiterativamente reportado en diferentes investigaciones, donde informan que la impulsividad no presenta diferencias de sexo (Barratt, 1994; Patton y Stanford, 2011; Patton *et al.*, 1995).

Existen estudios que reportan una estructura de dos factores, que se mencionan seguidamente. Es el caso de lo encontrado por Ireland y Archer (2008) con su muestra de internas penitenciarias mencionado anteriormente, en donde el modelo propuesto de tres factores que tuvo un buen ajuste para los hombres, no se comportó de igual forma en las mujeres ($RMSEA=.12$, $GFI=.92$, $CFI=.88$, $ECVI=.69$). Es por esto que eliminaron el factor *distractibilidad*, quedando dos factores, logrando así que los indicadores mejoraran sustancialmente ($RMSEA=.09$, $GFI=.96$, $CFI=.96$, $ECVI=.23$). Por tanto de acuerdo a estos indicadores de ajuste, para las mujeres convictas la impulsividad se explica mejor por aspectos comportamentales y de habilidades de planeación. Loyola (2011) con 216 mujeres peruanas convictas informó que la estructura factorial hallada en el análisis exploratorio no era la misma a la encontrada originalmente. Posteriormente el análisis confirmatorio así lo mostró ($RMSEA=.093$, $SRMR=.10$, $AGFI=.68$). Luego de haber eliminado algunos ítems poco significativos, probó una estructura de dos factores que arrojó mejores indicadores ($RMSEA=.058$, $SRMR=.078$) en la que no se incluía el factor *impulsividad atencional*. La ausencia de este factor fue referenciada asimismo por Fossati *et al.* (2001), Yang *et al.* (2007) y Haden y Shiva (2008). El primer factor fue designado como *impulsividad motora-atencional* compuesto en su mayoría por ítems de las escalas *impulsividad motora* e *impulsividad atencional*. El segundo, llamado *impulsividad no planeada-atencional* conformado por los ítems de *impulsividad no planeada* e *impulsividad atencional*. El factor *impulsividad motora-atencional* alcanzó una fiabilidad aceptable de $\alpha=.76$, mientras que *impulsividad no planeada-atencional* una de $\alpha=.69$ (Loyola, 2011). Dos factores también fueron informados por Haden y Shiva (2008) con pacientes hombres prisioneros, después que analizaron cinco diferentes modelos, el

que mejor se ajustaba a los datos era el conformado por una escala de *impulsividad motora* y otra de *impulsividad no planificadora*. La consistencia total para esta muestra fue de $\alpha=.71$.

1.6 Adaptaciones del test BIS-11 en niños y adolescentes

Frecuentemente las escalas autoadministradas para niños no son tan utilizadas como en adultos, porque se considera que en niños son menos precisas cuando miden su propio comportamiento. Pero por otra parte, existe evidencia que demuestra que los niños son mejores informantes que los padres y profesores (Fink y MacCown, 1993; Rapee, Barrett, Dadds y Evans, 1994). No obstante las anteriores consideraciones, el BIS exhibe al igual que en la adultez, una buena validez convergente en la niñez (Cosi, Hernández-Martínez, Canals y Vigil-Colet, 2011; Esteban y Taberero, 2011; Gilbert *et al.*, 2011; Morales-Vives, 2007; Nandagopal, Fleck, Adler, Mills, Strakowski y Del Bello, 2011).

En lo que tiene que ver con la población de niños y adolescentes, la escala ha tenido algunas adaptaciones que han mostrado, como lo ocurrido en adultos, la inconsistencia del modelo de tres factores de Patton *et al.* (1995). Veamos a continuación algunos de estudios más representativos. En primera instancia se tienen los que han informado acerca de una estructura de tres factores, como el de Von Diemen *et al.* (2007) en una adaptación para adolescentes brasileños con una muestra de 464 sujetos ($M=17.3$, $DT=1.7$). El análisis factorial exploratorio identificó tres factores, pero la solución fue diferente a la original, además siete de los ítems

tuvieron que ser retirados de la solución por presentar saturaciones menores de .30. El primer factor estuvo compuesto por ítems pertenecientes a la escala *impulsividad atencional* e *impulsividad no planificadora*. El segundo, por todos los ítems de la escala *impulsividad motora*, pero contenía algunos de la escala *impulsividad atencional* e *impulsividad no planificadora*. Y en el tercero, sólo ítems de la escala *impulsividad no planificadora*. La consistencia interna total reportada fue de $\alpha=.62$, que es una de las más bajas para el BIS-11 en las investigaciones revisadas (Vasconcelos *et al.*, 2012). Los autores consideran que la falta de ajuste del modelo y los bajos coeficientes de fiabilidad, pueden ser debidos al bajo nivel educativo de la muestra. Sin embargo, al revisar el proceso de adaptación de la versión inglesa a la brasilera, se observa que no se han seguido los delineamientos para la adaptación de test propuestos por Hambleton (2005) y Hambleton y Zenisky (2011), aspecto que pudo llevar a que se dieran sesgos indeseados, que muy seguramente afectaron la validez del cuestionario. Otra versión que sí ha podido replicar adecuadamente los tres factores es la que recientemente Recio (2012) realizó sobre el BIS-11 con una muestra aleatoria estratificada de 2082 niños estudiantes de 4º y 5º de primaria entre 9 y 12 años, y de 3º y 4º de secundaria entre 14 y 16 años, de la comunidad de Madrid. Partiendo de la versión española para adultos de Oquendo *et al.* (2001) modificó el enunciado de 19 de los 30 ítems para adaptarlo al rango de edad de los sujetos estudiados. Llevó a cabo un análisis factorial confirmatorio evaluando los índices de ajuste para tres modelos: unifactorial ($GFI=.95$, $NNFI=.90$, $CFI=.91$, $ECVI=2.04$, $RMSEA=.059$), bifactorial ($\Delta\chi^2=72,1$, $GFI=.95$, $NNFI=.91$, $CFI=.92$, $ECVI=1.96$, $RMSEA=.057$) y trifactorial ($\Delta\chi^2=27,8$, $GFI=.95$, $NNFI=.91$, $CFI=.92$, $ECVI=1.93$, $RMSEA=.057$). Y aunque los tres modelos poseen ajustes considerados como

adecuados de acuerdo a los índices globales calculados (*GFI*, *NNFI*, *CFI*, mayor de .90 y *RMSEA* cercano a .05), hay diferencias estadísticamente significativas en el incremento del χ^2 en los modelos propuestos, obteniendo la solución de tres factores el mejor ajuste en este indicador. Igualmente, al utilizar el índice de validación cruzada (*ECVI*) verificó que el valor menor lo obtiene el modelo trifactorial. De acuerdo a lo expuesto, la autora considera que de los tres, el que mejor ajuste exhibe es el de tres factores. En el análisis factorial, tres de los ítems fueron problemáticos y tuvieron bajas saturaciones (*Me muevo y ando más rápido que mis amigos*, *Estoy más interesado en el presente que en el futuro* y *Hago planes para cuando sea mayor*), quedando la versión final con 27 ítems. La consistencia interna total de la escala fue alta con $\alpha=.83$. Los coeficientes de las escalas fueron: para *impulsividad motora* de $\alpha=.67$ (9 ítems), *impulsividad no planificadora* $\alpha=.63$ (9 ítems) e *impulsividad atencional* $\alpha=.65$ (10 ítems) (Recio, 2012). Por último, está la versión de Cosi, Vigil-Colet, Canals (2008b) que ha identificado tres factores, será abordada en una sección posterior.

En lo que respecta a las versiones que han informado dos factores se tiene la de Hartmann, Rief y Hilbert (2011) quienes con una muestra de 659 adolescentes alemanes entre 10 y 20 años hicieron una adaptación del BIS-11. El análisis factorial exploratorio con rotación Promax exhibió tres factores, pero compuestos por ítems diferentes a los de la versión original de Patton *et al.* (1995). El primero incluía ítems de las escalas *impulsividad motora* e *impulsividad atencional*. El segundo, ítems de la escala de *impulsividad no planificadora*. Y el tercero, ítems de la escala *impulsividad motora*. La consistencia interna total fue de $\alpha=.74$. Yao *et al.* (2007) con una muestra

de 396 estudiantes de secundaria chinos, utilizaron el análisis factorial confirmatorio e informaron que el mayor ajuste a los datos fue para un modelo de seis factores de primer orden y dos de segundo orden. El primer factor contenía ítems de la escala *impulsividad motora* e *impulsividad atencional*, y el segundo, sólo ítems de *impulsividad no planificadora*. El coeficiente alfa Cronbach fue de $\alpha=.80$ y la correlación test-retest de $r=.70$.

Es de agregar, que existe una versión preliminar para adolescentes hindúes de Singh *et al.* (2008), de la que aún no se ha reportado la estructura factorial, ni la consistencia interna. Llama la atención el proceso que se ha llevado para adaptar la versión original inglesa para adultos a las características y actividades de estas edades, teniendo en cuenta aspectos culturales, socioeconómicos y estilos de vida. Lo que ha llevado a que varios de los ítems difieran de otras adaptaciones, como la italiana, según se explica en la publicación. No obstante como se mencionó anteriormente, falta conocer cómo serán las propiedades psicométricas de esta adaptación para poder realizar las comparaciones y análisis del caso.

En las secciones siguientes se revisará el proceso que se siguió para obtener la versión italiana para adolescentes, y la transformación que se llevó cabo a partir de ésta, para llegar a la adaptación española para niños (Cosi *et al.*, 2008b; Cosi, Vigil-Colet, Canals y Lorenzo-Seva, 2008c; Fossati, Barratt, Acquarini y Di Ceglie, 2002). Estas dos versiones, en especial la última, fueron fundamentales en el proceso que se ha llevado para obtener la adaptación colombiana.

1.6.1 La adaptación italiana para adolescentes: BIS-11A

La versión italiana denominada *Barratt Impulsiveness Scale-11 for Adolescents* (en adelante BIS-11A) tuvo en cuenta las dificultades que se presentaban con la comprensión de los ítems por parte de los adolescentes ($M=16.4$, $DT=1.5$) con la versión para adultos. Para esto, los tres autores italianos, los mismos que la habían adaptado de la versión inglesa a la muestra adulta, modificaron los ítems con el objeto de lograr una mejor comprensión y contextualización, y tener una adecuada correspondencia a las experiencias y conductas de los adolescentes. En este proceso, a algunos se les modificó parte su redacción, como es el caso del ítem 1 *I plan tasks carefully* que fue cambiado por *I plan what I have to do*, el ítem 6 *I have "racing" thoughts* cambiado por *My thoughts are racing too fast*, el ítem 11 *I "squirm" at plays or lectures* cambiado por *I cannot stand still at movies or school*, el ítem 25 *I spend or charge more than I earn* cambiado por *I spend more than I should*, el ítem 28 *I am restless at the theater or lectures* cambiado por *I am restless at the movies or lectures* y el ítem 29 *I like puzzles* cambiado por *I like to play chess or checkers*. Por otra parte, cuatro ítems que tenían en su redacción unas características muy relacionadas con la vida y manera de pensar adulta, fueron reescritos, el ítem 16 *I plan trips well ahead of time* quedó como *I plan my spare time*, el ítem 16 *I change jobs* quedó como *I change my mind about what I will do when I grow up*, el ítem 21 *I change residences* quedó como *I change friends* y el ítem 26 *I often have extraneous thoughts when thinking* quedó como *When I think about something, other thoughts pop up in my mind*. En definitiva, la mitad de los ítems fueron modificados parcial o totalmente (Fossati *et al.*, 2002).

En lo correspondiente al análisis factorial, llevaron a cabo siete análisis confirmatorios (i.e., nulo, unidimensional, tres factores correlacionados, seis factores correlacionados, seis factores correlacionados de primer orden con un factor de segundo orden, seis factores correlacionados de primer orden con tres factores de primer orden, y finalmente, seis factores correlacionados de primer orden con dos de segundo orden). Los modelos que obtuvieron mejores ajustes fueron tres. El primero, de seis factores de primer orden y uno de segundo orden, con índices de ajuste de $RMSEA=.017$, $NNIF=.98$, $CFI=.98$ y $\chi^2/gl=1.16$. El segundo, de seis factores de primer orden y tres de segundo orden con $RMSEA=.015$, $NNIF=.98$, $CFI=.98$ y $\chi^2/gl=1.12$. Y el tercero, de seis factores de primer orden y dos de segundo orden con $RMSEA=.010$, $NNIF=.99$, $CFI=.99$ y $\chi^2/gl=.90$, obtuvo un ajuste superior, ya que todos los indicadores de bondad y ajuste así lo informaban, y fue el que finalmente se seleccionó. Si bien lo propuesto por los autores, hay que anotar en este punto, que si se observa con detenimiento, el modelo de seis factores de primer orden y tres de segundo orden, que fue descartado, tuvo un buen ajuste al modelo trifactorial. En el modelo seleccionado, el primer factor de segundo orden denominado *impulsividad general*, incluía los factores de primer orden *impulsividad motora*, *atención*, *perseverancia* y *falta de demora en la obtención de la gratificación inmediata*. Es decir una mezcla de los factores *impulsividad motora* e *impulsividad atencional* propuestos por Patton, *et al.* (1995). El segundo factor de segundo orden designado como *impulsividad no planificadora*, incluía los factores de primer orden *autocontrol* e *impulsividad no planificadora*. La correlación entre los dos factores fue de $r=.53$. Los ítems que presentaron cargas factoriales más bajas fueron el ítem 21 *I change friends*, el ítem 23 *I can think about one problem at time*, el ítem 24 *I change hobbies*

and sports y el ítem 29 *I like to play chess or checkers*, que igualmente exhibieron los valores más bajos en la versión para adultos. En el análisis de los ítems, también los mismos ítems tuvieron bajas correlaciones ítem total ($<.1$). El instrumento mostró una consistencia interna satisfactoria de $\alpha=.78$, los demás coeficientes no fueron informados en la publicación (Fossati *et al.*, 2002; Fossati *et al.*, 2001; Patton *et al.*, 1995).

Según lo anterior, y ante la identificación de un modelo de dos factores de segundo orden, los autores señalaron que a diferencia de la adultez, en la adolescencia los componentes cognitivos, motores y atencionales de la impulsividad, no se encuentran bien diferenciados. Además, la impulsividad tendría que ser conceptualizada de una manera bidimensional más que tridimensional. Por lo tanto, lo más apropiado en estas edades sería interpretar las puntuaciones de la escala total, en lugar de centrarse en las puntuaciones de las subescalas (es decir, los diferentes factores) que componen el BIS-11A (Fossati *et al.*, 2002; Patton *et al.*, 1995).

1.6.2 La adaptación española para niños: BIS-11c

Cosi *et al.* (2008c) cuestionan el procedimiento utilizado en la versión italiana para adolescentes de Fossati *et al.* (2002) debido a la inadecuación de la metodología empleada para determinar los factores. Ellos consideran que el procedimiento utilizado en su desarrollo no es el más apropiado cuando se está desarrollando un nuevo instrumento, como en el caso del BIS-11A, ya que las características del constructo de impulsividad varían entre la adolescencia y la adultez, y un análisis de

tipo confirmatorio no es el más adecuado para este caso. Plantean que en esta etapa de desarrollo de la escala, es mucho más apropiado realizar un análisis factorial exploratorio que uno confirmatorio. Este tipo de análisis permitiría identificar y eliminar, si es el caso, los ítems que sean problemáticos. Al respecto señalan por ejemplo que la adaptación italiana posee cuatro ítems con una bajas cargas factoriales ($<.16$). Y además llaman la atención respecto a que no se suministra información completa de todos los coeficientes de fiabilidad, ya que sólo se informa la fiabilidad total del test, quedando sin reportar la consistencia de los factores de primer y segundo orden.

Esta adaptación española para niños llamada *Barratt Impulsiveness Scale-11 for children* (en adelante BIS-11c), fue desarrollada a partir de los ítems de la versión italiana, siguiendo las orientaciones propuestas para estos caso por Hambleton (2005). Una vez que se comprobó que no había diferencias importantes entre la traducción y la retrotraducción de los ítems, y estos eran adecuados para su aplicación en niños españoles, se hizo un primer estudio con de 377 sujetos con edades comprendidas entre los 8 y 12 años ($M=10.4$, $DT=.90$). Utilizando el programa Factor (Lorenzo-Seva y Ferrando, 2006) para llevar a cabo el análisis factorial exploratorio. A diferencia de lo reportado para la muestra italiana, tres factores fueron identificados: *impulsividad motora*, *impulsividad no planificadora* e *impulsividad cognitiva*. Pero dos ítems tuvieron que ser eliminados por tener comunalidades inferiores a $.06$. Estos fueron el ítem 27 (*I am more interested in the present than the future*), posiblemente por su dificultad de comprensión para los niños y el ítem 29 (*I like to play chess o checker*), porque se refiere a juegos que no son frecuentemente realizados por los niños

españoles en la actualidad. Los coeficientes de confiabilidad obtenidos con esta versión preliminar de 28 ítems fueron buenos para escala total con $\alpha=.76$, *impulsividad motora* con $\alpha=.82$ e *impulsividad no planificadora* con $\alpha=.71$. Más *impulsividad cognitiva* con tan sólo tres ítems (*I make up my mind quickly*, *My thoughts are racing too fast* y *I concentrate easily*) tuvo una consistencia baja con $\alpha=.60$ (Cosi *et al.*, 2008c).

Para mejorar los coeficientes alfa Cronbach de la escala *impulsividad cognitiva* se agregaron cinco nuevos ítems que fueron redactados sobre la base la escala *impulsividad funcional* para niños de Brunas-Walgstaff, Tilley, Verity, Ford y Thomson (1997).

Se administró un nuevo cuestionario de 33 ítems a una muestra de 456 niños entre 9 y 13 años ($M=11.0$, $DT=.90$). Nuevamente los datos fueron analizados, lo que llevó a la eliminación de siete ítems, seis por poseer bajas cargas factoriales, y uno por tener un contenido prácticamente idéntico a otro. De los que tuvieron bajas saturaciones, el ítem 4 (*Soy muy despreocupado*), el ítem 19 (*Me gusta pensar en problemas complicados*) y el ítem 29 (*Cambio frecuentemente de deportes y aficiones*) presentaron igualmente bajas cargas en Fossati *et al.* (2002). En resumen, en la versión definitiva terminaron por eliminarse seis ítems de la versión italiana, y se agregaron a la escala *impulsividad cognitiva* dos nuevos provenientes de la escala *impulsividad funcional* (*Cuando mis amigos me preguntan algo, puedo responder rápidamente* y *Soy de los primeros en levantar la mano en clase cuando el profesor hace una pregunta*) (Brunas-Walgstaff *et al.*, 1997; Cosi, 2008; Cosi *et al.*, 2008b).

Se identificaron nuevamente en esta versión final los tres factores: *impulsividad motora* (13 ítems), *impulsividad no planificadora* (8 ítems) e *impulsividad cognitiva* (5 ítems), que correspondían a la conceptualización teórica de Barratt (1985). No obstante hay que aclarar que aunque la solución factorial es la misma, hay dos ítems que no pertenecen a la versión original, como ya se ha visto, y el contenido de gran parte de los ítems no es exactamente el mismo que el de la original inglesa para adultos. Esto debido a los cambios y adiciones que han tenido que realizarse para adaptarlo a las características de los niños, primero italianos y posteriormente españoles. Los ítems incluidos en la escala *impulsividad motora* se encontraban relacionados con los factores de primer orden descritos por Fossati *et al.* (2002) (es decir, *impulsividad motora*, *atención*, *falta de demora* y *perseverancia*) y que correspondían al factor *impulsividad general*. Los ítems incluidos en la escala *impulsividad no planificadora*, estuvieron relacionados con *autocontrol* y *complejidad cognitiva*, y correspondían al factor *impulsividad no planificadora* de la versión italiana. Pero, los ítems de la escala *impulsividad cognitiva* no estaban relacionados con ninguno de los dos factores propuestos en la versión italiana. En cambio, este factor se encuentra muy asociado al concepto de *impulsividad funcional* reportada para niños (Cosi, Morales-Vives, Canals, Lorenzo-Seva y Vigil-Colet, 2008a; Brunas-Walgstaff *et al.*, 1997). De hecho, esta rapidez en el procesamiento de la información y en la toma de decisiones, concuerda con la *impulsividad cognitiva* de Barratt (1985). Los indicadores de bondad de ajuste al modelo del BIS-11c fueron buenos ($RMSEA=.043$, $CFI=.91$). Posee una fiabilidad de $\alpha=.80$ para *impulsividad motora*, $\alpha=.71$, para *impulsividad no planificadora* y $\alpha=.68$ para *impulsividad cognitiva*. La

correlación ítem total de los 26 ítems estuvo entre $r=.20$ y $r=.80$ (Cosi, 2008; Cosi *et al.*, 2008b; Cosi *et al.*, 2008c).

Por el procedimiento seguido y las propiedades psicométricas informadas, tanto a nivel de consistencia interna, como de réplica de la estructura factorial original, se ha tomado como base el BIS-11c para realizar la adaptación en niños colombianos.

2 El Cuestionario de Agresividad de Buss y Perry

2.1 Conceptualización de agresividad

En el lenguaje común la palabra *agresividad* es definida como la tendencia que tiene un individuo a actuar o a responder violentamente, y la palabra *agresión*, como el acto de acometer a otro para matarlo, herirlo o hacerle daño (Real Academia Española, 2001). Si bien existen múltiples definiciones de agresividad, según la perspectiva teórica ésta ha tenido diversas conceptualizaciones. Bandura (1983) la concibe como una conducta perjudicial y destructiva que viola las normas sociales. Berkowitz (1996) como un comportamiento dirigido, que tiene la intención de herir física o psicológicamente a otro. Y Barratt (1991) en su conceptualización la clasifica

en tres categorías: agresividad premeditada o adquirida, agresividad por motivos médicos y agresividad impulsiva.

Y aunque hay un acuerdo referente al comportamiento agresivo, no ha ocurrido lo mismo con sus diversos componentes, algunos estudios los han utilizado indistintamente, mientras que otros han buscado diferenciarlos. Esta falta de claridad debida al solapamiento entre las diferentes teorías, ha llevado a un desacuerdo entre los distintos autores (Barratt y Slaughter, 1998; Suris, Lind, Emmetta, Bormanc, Kashner y Barratt, 2004). Ni tan sólo desde la psiquiatría hay una definición general aceptada de agresividad (American Psychiatric Association, 2000).

Para los efectos de esta tesis se ha tomado el abordaje multidimensional de agresividad propuesto por Buss y Perry (1992), que la consideran como un constructo compuesto por tres niveles bien diferenciados e interrelacionados: conductual (físico y verbal), emocional y cognitivo.

2.2 Los orígenes: The Hostility Inventory de Buss y Durkee

El cuestionario desarrollado por A. H. Buss y M. P. Perry tiene sus comienzos en el *The Hostility Inventory* (en adelante BDHI) desarrollado en una época en la que no existían medidas para evaluar los componentes de la agresividad. Fue diseñado para medir las diferencias individuales en el rasgo hostilidad, y una sus características más importantes era su carácter multidimensional. Posee 75 ítems con respuestas verdadero-falso, de los cuales 66 ítems miden hostilidad y 9 ítems culpa. El

instrumento basado en consideraciones teóricas posee siete escalas: *ataque* (violencia hacia los otros), *hostilidad indirecta* (hablar con rodeos y agresión indirecta), *negativismo* (comportamiento oposicional usualmente hacia la autoridad), *irritabilidad* (estar propenso a explotar a la menor provocación), resentimiento (celos y odio hacia los otros), *desconfianza* (proyectar la hostilidad hacia los demás) y *hostilidad verbal* (afecto negativo expresado cuando se habla, tanto en la forma como en el contenido). Estas dimensiones permitían conocer no sólo si una persona era agresiva, sino lo más importante, la forma en que la agresión era manifestada (Buss y Durkee, 1957; Buss y Perry, 1992).

El análisis del BDHI identificó dos factores generales, uno que involucraba un *componente motor*, relacionado con hostilidad abierta, compuesto por *ataque*, *hostilidad indirecta*, *irritabilidad* y *hostilidad verbal*. El otro, denominado *componente emocional*, implicaba los aspectos experienciales e incluía *resentimiento* y *desconfianza* (Buss y Durkee, 1957). Estos dos factores fueron hallados posteriormente en otras investigaciones y se les llamó: *hostilidad neurótica* y *hostilidad expresiva*. El componente *hostilidad neurótica* tendía a incluir las subescalas *resentimiento* y *desconfianza*. Pero el de *hostilidad expresiva*, a excepción de las subescalas *ataque* y *hostilidad indirecta*, variaba entre las muestras y era inestable, como ocurría con *hostilidad indirecta* e *irritabilidad* (Bushman, Cooper, y Lemke, 1991; Buss y Durkee, 1957; Felsten y Leitten, 1993; Siegman, Dembroski y Ringel, 1987). En cuanto su constructo de validez, existe una amplia evidencia para el total de la escala, y presenta correlaciones significativas con diversas medidas

psicológicas (Bishop y Quah 1998; Buss y Perry 1992; Gunn y Gristwood, 1975; Matthews y Saal, 1978; Petzel y Michaels, 1973).

No obstante la amplia popularidad del BDHI, éste presentaba varias dificultades en lo relacionado a sus características psicométricas (Biaggio, Supplee y Curtis, 1981). Aunque el análisis encontró dos factores generales que se relacionaban con el aspecto motor y emocional, las subescalas fueron establecidas a priori, y los ítems asignados a cada una de ellas según su aparente validez. Pero lo más importante, no existía un análisis factorial de los ítems que demostrara a qué factor correspondía cada una (Bendig, 1962; Edmunds y Kendrick, 1980). La anterior situación fue asociada a varios problemas. El primero, la falta de estabilidad de la escala en el tiempo, ya que no se había presentado evidencia sobre la fiabilidad test-retest. El segundo, el tipo de respuesta dicotómica utilizado (verdadero-falso), no era el más adecuado, ya que limitaba sensiblemente las opciones de respuesta y afectaba a las estimaciones. Quienes respondieron este cuestionario tenían dificultades con el formato de respuesta y señalaron que preferirán posibilidades de respuesta más amplias. En realidad, las conductas evaluadas se adaptaban más a una escala continua, como es el caso de la Likert con cinco opciones, que fuera desde el menor hasta el mayor valor. El tercero, algunos ítems tenían cargas factoriales en más de una subescala y se solapaban. Por ejemplo, un mismo ítem podía pertenecer a dos subescalas, tales como *agresión indirecta* y *agresión verbal* (Buss y Perry, 1992).

2.3 *Desarrollo del test AQ*

Este cuestionario se construyó teniendo como base los ítems del BDHI, utilizando una muestra de 1253 estudiantes universitarios. Para esto se tomaron algunos de sus ítems sin ningún tipo de variación, otros tuvieron que ser cambiados en su redacción para alcanzar una mayor claridad y comprensión, varios fueron removidos por ser considerados ambiguos o problemáticos, y finalmente, se agregaron nuevos ítems. Como resultado se obtuvo un grupo inicial de 52 ítems para comenzar a realizar el análisis. Se buscaba con estos ítems evaluar seis dimensiones iniciales para la agresión: *agresividad física*, *agresividad verbal*, *ira*, *agresión indirecta*, *resentimiento* y *desconfianza*. Se llevó a cabo entonces con una primera muestra un análisis factorial exploratorio de ejes principales con rotación oblicua, pues se esperaba que varios de los componentes pudieran estar correlacionados. El análisis arrojó cuatro factores que fueron denominados como: *agresividad física*, *agresividad verbal*, *ira* y *hostilidad*. Este último contenía ítems de *resentimiento* y *desconfianza*. Sobre esta solución se seleccionaron sólo aquellos ítems con cargas factoriales superiores o iguales a .35 en su respectivo factor, igualmente cada ítem debía tener cargas inferiores a este mismo valor en los demás factores. Lo anterior llevó a eliminar 23 ítems, quedando finalmente sólo 29 ítems. Luego, con las otras dos muestras restantes se hicieron sendos análisis confirmatorios. El modelo que mejor ajuste presentó fue el de cuatro factores, además, éste encuadraba en su planteamiento teórico sobre la agresividad (Buss y Perry, 1992).

La versión definitiva fue denominada *The Buss-Perry Aggression Questionnaire* (en adelante AQ), compuesto por nueve ítems que corresponden a *agresividad física*, cinco ítems a *agresividad verbal*, siete ítems a *ira* y ocho ítems a *hostilidad*. Los factores *agresividad física* y *agresividad verbal*, asociadas con herir o hacer daño a otros, representan el componente instrumental. El factor *ira*, que implica una agitación fisiológica y una preparación para la agresión, representa el componente emocional o afectivo. Y finalmente, el factor *hostilidad*, constituido por sentimientos de malestar e injusticia, representa el componente cognitivo. El formato para responder varió de verdadero-falso a una escala tipo Likert, que va de uno (extremadamente incharacterístico en mí) a cinco (extremadamente característico en mí) (Buss y Perry, 1992).

La consistencia interna para el total del cuestionario fue de $\alpha=.89$, para *agresividad física* de $\alpha=.85$, *agresividad verbal* de $\alpha=.72$, *ira* de $\alpha=.83$ y *hostilidad* de $\alpha=.77$. Adicionalmente, mostró una adecuada estabilidad en el tiempo, las correlaciones test-retest fueron para el total de $r=.89$, *agresividad física* de $r=.80$, *agresividad verbal* de $r=.76$, *ira* de $r=.72$ y *hostilidad* de $r=.72$. La correlación entre los factores para toda la muestra fue bastante buena, a excepción de los valores entre *hostilidad* con *agresividad física* y *agresividad verbal*. De este modo, *agresividad física* correlacionó con *agresividad verbal* ($r=.45$), con *ira* ($r=.48$) y con *hostilidad* ($r=.28$). La *agresividad verbal* con *ira* ($r=.48$) y con *hostilidad* ($r=.25$), e *ira* con *hostilidad* ($r=.45$) (Buss y Perry, 1992). Como se observa, el factor *ira* correlacionó fuertemente con todos los factores, resultado que sugiere que éste puede ser una especie de puente entre los componentes instrumentales y cognitivos de la agresión,

aspecto que ha sido señalado posteriormente por otros autores (Andreu, Peña y Graña, 2002; Buss y Perry 1992; Meesters, Muris, Bosma, Schouten y Beuving, 1996). En lo relacionado a las diferencias de género, los hombres obtuvieron puntajes más altos en todas las escalas, a excepción de *hostilidad*, más las diferencias mayores se dieron en *agresividad física*, que ha sido evidenciado en distintos estudios (Ang, 2007; Fossati, Maffei, Acquarini y Di Ceglie, 2003; Nakano, 2001; Reyna, Lello, Sanchez y Brussino, 2011; Santisteban, Alvarado y Recio, 2007).

2.4 Transculturalidad del test AQ

El AQ ha tenido un gran número de adaptaciones a otras culturas e idiomas de muy diversas características. Por ejemplo, la versión original ha sido validada en muestras de población norteamericana y canadiense (Bernstein y Gesn, 1997; Harris, 1995, 1997), británica (Archer, Holloway y McLoughlin, 1995a; Archer, Kilpatrick y Bramwell, 1995b), sueca (Prochazka y Agren, 2001), holandesa (Meesters *et al.*, 1996), japonesa (Nakano, 2001), italiana (Fossati *et al.*, 2003), española (García-León, Reyes, Vila, Pérez, Robles y Ramos, 2002), china (Maxwell, 2007), griega (Tsorbatzoudis, 2006) y alemana (Von Collani y Werner, 2005). También existen versiones para adolescentes (Reyna *et al.*, 2011; Santisteban y Alvarado, 2009). E inclusive, ha sido adaptado a poblaciones con problemas de delincuencia, como las reportadas por Morren y Meesters (2002) y Williams, Boyd, Cascardi y Poythress (1996).

No obstante, además de la versión de 29 ítems, existen otras refinadas que han sido validadas en distintas culturas, como es el caso de la desarrollada por Bryant y Smith (2001) con población británica, norteamericana y canadiense. Y a partir de ésta, se efectuaron adaptaciones a otros idiomas, entre los que se cuenta el chino (Ang, 2007), griego (Vitoratou, Ntzoufras, Smyrnis, y Stefanis, 2009) y español (Gallardo-Pujol, Kramp, García-Forero, Pérez-Ramírez y Andrés-Pueyo, 2006). Y en esta línea de las versiones cortas, se encuentran aquellas que buscan encontrar una estabilidad entre las distintas culturas, utilizando el menor número de ítems posible, como la realizada por Vigil-Colet, Lorenzo-Seva, Codorniu-Raga y Morales-Vives (2005).

2.5 El modelo del test AQ en adultos

Los estudios realizados a nivel exploratorio y confirmatorio en adultos no siempre han replicado la estructura factorial inicial del instrumento. Algunas veces el modelo encontrado es adecuado, en otras, los ítems no saturan en las escalas originales. Es por esto que se analizarán a continuación las diversas soluciones factoriales encontradas.

2.5.1 Estudios que informan del modelo tetrafactorial

Gran parte de las investigaciones realizadas después de la publicación original han podido replicar los factores propuestos por Buss y Perry (1992). Como es el caso de Fossati *et al.* (2003) quienes se centraron en la replicabilidad del modelo, concluyendo que éste es consistente en adultos y en estudiantes de secundaria. El

coeficiente de fiabilidad hallado para la escala *agresividad física* fue de $\alpha=.85$, *agresividad verbal* de $\alpha=.53$, *ira* de $\alpha=.72$, *hostilidad* de $\alpha=.78$ y para la escala total en hombres de $\alpha=.89$ y en mujeres de $\alpha=.87$.

Por otro lado, hay los que identifican los cuatro factores con algunas objeciones respecto a los ítems y al nivel del ajuste. Nakano (2001) con estudiantes de psicología llevó a cabo un análisis factorial exploratorio de ejes principales y rotación Oblimin. Los resultados soportaron el modelo original. Pese a esto, en la revisión de las cargas factoriales encontró dos ítems con bajas saturaciones, el ítem 7 (*I can think of no good reason for ever hitting a person*) perteneciente al factor *agresividad física* (.30) y el ítem 4 (*I am an even-tempered person*) correspondiente al factor *ira* (.09). Estos dos ítems adicionalmente tenían la particularidad de ser invertidos y estaban negativamente orientados, por lo que se propuso que se retiraran para estudios transculturales (Mook, Kleijn y Van der Ploeg, 1991; Nakano, 2001). La eliminación de estos dos ítems también ha sido planteada por Porras, Salamero y Sender (2001) y Sierra y Gutiérrez (2007). Los factores no obstante tuvieron una aceptable consistencia interna para *agresividad física* con $\alpha=.77$, *agresividad verbal* con $\alpha=.77$, *ira* con $\alpha=.76$ y *hostilidad* con $\alpha=.75$ (Nakano, 2001).

Andreu *et al.* (2002) hicieron una adaptación del AQ con 29 ítems en adultos jóvenes. En primer lugar, realizaron un análisis factorial exploratorio aplicando el método de componentes principales y rotación Varimax, que arrojó una estructura de cuatro factores. Pese a lo anterior, tres de los ítems tuvieron cargas factoriales por encima de .35 en más de un factor. Específicamente, el ítem 18 (*Mis amigos dicen que*

discuto mucho), perteneciente a escala *agresividad verbal*, saturó en *ira* (.53), el ítem 22 (*Algunas veces pierdo los estribos sin razón*), perteneciente a la escala *ira*, saturó en *hostilidad* (.40) y el ítem 11 (*Algunas veces me siento tan enfadado como si estuviera a punto de estallar*), perteneciente a la escala *ira*, saturó en *hostilidad* (.46). Posteriormente, llevaron a cabo un análisis factorial confirmatorio, encontrando que todos los ítems obtuvieron altos coeficientes de regresión estandarizados, exceptuando tres de ellos que fueron poco discriminativos. Estos fueron: el ítem 23 (*Desconfío de desconocidos demasiado amigables*) de la escala *hostilidad*, el ítem 29 (*He llegado a estar tan furioso que rompía cosas*) de la escala *agresividad física*, y el ítem 19 (*Algunos de mis amigos piensan que soy una persona impulsiva*) de la escala *ira*. Con referencia al ajuste de los indicadores absolutos y relativos, éstos fueron adecuados, lo que significa que el modelo fue apropiadamente replicado ($RMSR=.05$, $GFI=.93$, $AGFI=.92$, $RMSEA=.05$). Finalmente, los coeficientes de fiabilidad fueron buenos tanto para la escala total ($\alpha=.88$) como para *agresividad física* ($\alpha=.86$) e *ira* ($\alpha=.77$). Pero por otro lado, las escalas *agresividad verbal* y *hostilidad* tuvieron valores menores ($\alpha=.68$ y $\alpha=.72$).

Harris (1995) llevando a cabo un análisis factorial con máxima verosimilitud, en una muestra de estudiantes universitarios canadienses, señaló que su solución tuvo un moderado ajuste al modelo original ($GFI=.825$, $AGFI=.794$, $RMSR=.077$). No obstante, al revisar las cargas factoriales del ítem 6 (*I am suspicious of overly friendly strangers*) y del ítem 8 (*When people are especially nice, I wonder what they want*) del factor *hostilidad*, relacionados con *desconfianza*, vio que presentaban valores bajos en comparación con los otros (saturaciones de .35 y .37

respectivamente), y consideró que éstos evaluaban más *desconfianza* que *hostilidad*. Propuso eliminarlos, quedando el factor *hostilidad* con seis ítems, además, obtenía un pequeño incremento del coeficiente alfa Cronbach de $\alpha=.799$ a $\alpha=.811$. Por lo demás, la consistencia interna de las escala *agresividad física*, *agresividad verbal* e *ira*, arrojaron unos valores similares a los encontrados en la versión original del AQ ($\alpha=.84$, $\alpha=.75$, $\alpha=.83$). Posteriormente, Harris (1997) en un nuevo estudio utilizando todos los ítems, reportó coeficientes para *agresividad física*, *agresividad verbal*, *ira* y *hostilidad* de $\alpha=.75$, $\alpha=.70$, $\alpha=.80$ y $\alpha=.82$.

Sierra y Gutiérrez (2007) en el contexto suramericano, específicamente en el Salvador, aplicando un análisis factorial exploratorio de componentes principales, encontraron los mismos factores de la estructura original, si bien eliminaron dos de los ítems: el ítem 3 (*Me enfado rápidamente pero se me pasa enseguida*) y ítem 24 (*No encuentro ninguna buena razón para pegarle a una persona*). La consistencia interna para el total de la muestra fue de $\alpha=.89$, y para *agresividad física* de $\alpha=.80$, *agresividad verbal* de $\alpha=.73$, *ira* de $\alpha=.72$ y *hostilidad* de $\alpha=.76$.

El modelo de cuatro factores fue igualmente informado por Meesters *et al.* (1996) con una muestra holandesa, realizaron el análisis confirmatorio obteniendo un moderado ajuste al modelo ($GFI=.87$, $AGFI=.84$, $RMSR=.064$). Aquí ellos estuvieron de acuerdo en que estos índices podían incrementarse al retirar dos ítems del factor *hostilidad* (*I am suspicious of overly friendly* y *When people are especially nice, I wonder what they want*). Al mismo tiempo propusieron que el modelo podía ser mejorado si se retiraba un ítem del factor *agresividad verbal* (*I tell my friends openly*

when I disagree with them). Al eliminar estos tres ítems, los indicadores fueron superiores ($GFI=.90$, $AGFI=.88$, $RMSR=.058$). La consistencia interna para la muestra total fue de $\alpha=.84$, para *agresividad física* de $\alpha=.75$, *agresividad verbal* de $\alpha=.50$, *ira* de $\alpha=.70$ y *hostilidad* de $\alpha=.76$. La eliminación de los dos ítems de la escala *hostilidad* ha sido cuestionada por Bernstein y Gesn (1997), quienes han señalado que estos ítems tienen puntajes más altos que los demás de la escala, y por tanto, consideran que agregan información al incrementar el rango del constructo *hostilidad*.

Finalmente, se han identificado cuatro factores con características diferentes a los originales. Como se lee en el artículo de García-León *et al.* (2002), que en un análisis factorial exploratorio de componentes principales con rotación Oblimin identificaron cuatro factores, pero con una estructura factorial distinta a propuesta por Buss y Perry (1992). Éstos fueron denominados: *ira con resentimiento*, *agresividad verbal*, *agresividad física* y *desconfianza*. En este modelo, el componente conductual no varió mucho en comparación con la escala original, más se observó que *agresividad física* tuvo un menor número de ítems y *agresividad verbal* uno mayor. En donde sí hubo diferencias fue en los componentes cognitivo y emocional. Hipotetizando en este sentido que los aspectos relacionados con *resentimiento* se encuentran más asociados al contenido emocional, mientras que los de *desconfianza* son más del componente cognitivo. El alfa Cronbach fue de $\alpha=.82$ para el total de la escala, para *agresividad verbal* de $\alpha=.63$, *agresividad física* de $\alpha=.63$, *ira con resentimiento* de $\alpha=.77$ y *desconfianza* de $\alpha=.67$.

2.5.2 Estudios que informan diferentes factores

Algunas publicaciones han señalado acerca de la inadecuación del modelo de Buss y Perry (1992), y han reportado factores distintos. Williams *et al.* (1996) en un análisis confirmatorio hallaron que el modelo de cuatro factores presentaba un ajuste pobre ($NNFI=.73$, $CFI=.84$). En cambio, propusieron basados en un nuevo análisis exploratorio, un modelo de dos factores. El primero consistente en un factor que combinaba *agresividad física e ira*, y el segundo, compuesto por *agresividad verbal y hostilidad*. Pero Diamond, Wang y Buffington-Vollum (2005) verificaron la estructura propuesta por Williams *et al.* (1996), y concluyeron que la estructura original de Buss y Perry (1992) ofrecía un mejor ajuste. Por su lado Fossati *et al.* (2003) también la rechazaron, pues vieron un inadecuado ajuste en este modelo bifactorial, y para completar, Bryant y Smith (2001) igualmente hallaron un deficiente ajuste en sus tres muestras.

En el sur de Italia, Sommantico, Osorio, Parrello, De Rosa y Donizzetti (2008) a través del análisis factorial exploratorio y confirmatorio, propusieron un modelo de tres dimensiones denominadas *agresividad física, hostilidad, e incapacidad para verbalizar la ira*, en donde las dos primeras corresponden a las escalas originales. El ajuste a este modelo de tres factores con 23 ítems fue $RMSEA=.06$, $NNFI=.82$, $GFI=.91$ y $AGFI=.89$. La consistencia interna para la escala total fue de $\alpha=.82$, y para cada una de las escalas de $\alpha=.78$, $\alpha=.67$ y $\alpha=.77$.

Y en Colombia Castrillón, Ortiz y Vieco (2005) con una muestra de 717 estudiantes universitarios con edades comprendidas entre 16 y 25 años, utilizaron un análisis exploratorio de componentes principales con rotación Varimax. Partiendo de una versión poco conocida del AQ con 40 ítems, propusieron un instrumento de 19 ítems con cinco factores. Éstos fueron: *déficit de autocontrol de la agresión física* ($\alpha=.81$), *percepción de hostilidad externa* ($\alpha=.86$), *déficit de autocontrol de la agresión verbal* ($\alpha=.80$), *desconfianza* ($\alpha=.57$) y *no agresión* ($\alpha=.41$). El coeficiente alfa Cronbach total estuvo en $\alpha=.82$. Los autores explican que esta estructura factorial puede ser debida a diferencias de tipo cultural entre los países anglosajones y nórdicos, con los latinoamericanos.

2.5.3 Las versiones refinadas del test AQ

También hay versiones reducidas que frecuentemente son utilizadas, una de las que más validez posee es la de Bryant y Smith (2001) quienes trabajaron con cinco muestras independientes, dos perteneciente a estudios anteriores, una británica de Archer *et al.* (1995a) y otra canadiense de Harris (1995), y tres exclusivamente conseguidas para su investigación. Al analizar el ajuste del modelo de cuatro factores con 29 ítems de Buss y Perry (1992), éste fue pobre en sus tres muestras, tanto en los indicadores de ajuste absoluto con *GFI* de .81, .76 y .81 y *RMSEA* de .077, .084 y .072, como en los de ajuste relativo con *CFI* de .81, .78 y .82 y *NNFI* de .79, .70 y .80. Como se ve, los índices de ajuste *GFI* y *CFI* estuvieron por debajo de .90. y fueron considerados inaceptables, además, los *RMSEA* estuvieron cercanos a .08, límite máximo para considerar un ajuste razonable. Estos datos fueron consistentes con

investigaciones anteriores que informaban de la inadecuación del modelo con 29 ítems (Archer *et al.*, 1995b; Harris, 1995; Williams *et al.*, 1996). De acuerdo a lo anterior, se consideró que la solución factorial obtenida con los 29 ítems no encajaba suficientemente con el marco conceptual planteado por Buss y Perry (1992).

Entonces, para preservar los fundamentos teóricos del modelo original y mejorar los índices de ajuste, se desarrolló una versión refinada. Para lo que se tuvieron que realizar varios pasos. Primero, eliminar los ítems que tuvieran una carga factorial menor a .40, con el objeto de incrementar la proporción de varianza explicada. Segundo, para dar una mayor claridad al modelo, se excluyeron los ítems que cargaban simultáneamente en dos o más factores con saturaciones superiores a .40. Tercero, para dar una mayor precisión, se descartaron los ítems que no reflejaban directamente los rasgos agresivos, éste fue el caso de los dos ítems invertidos, uno de ellos perteneciente a la escala *agresividad física*, y el otro a la escala *ira*. Como resultado se obtuvo una versión de 12 ítems. El ajuste mejoró sensiblemente, como lo exhiben los indicadores resultantes de ajuste absoluto *GFI* (.94, .93, .94) y *RMSEA* (.063, .068, .071), así como los de ajuste relativo *CFI* (.96, .95, .91) y *NNFI* (.94, .93, .87). Como se observa, hay una aceptable proporción de varianza común explicada tanto en los indicadores absolutos como en los relativos, que refleja un razonable ajuste al modelo de Bryant y Smith (2001). Posteriormente con las muestras cuatro y cinco se hallaron igualmente buenos indicadores para esta versión refinada, *GFI* (.90, .90), *RMSEA* (.092, .094), *CFI* (.92, .91) y *NNFI* (.90, .88). La fiabilidad de las escalas estuvo dentro de los parámetros: *agresividad física* ($\alpha=.79$, $\alpha=.80$), *agresividad verbal* ($\alpha=.83$, $\alpha=.80$), *ira* ($\alpha=.76$, $\alpha=.76$) y *hostilidad* ($\alpha=.75$, $\alpha=.70$).

Esta versión reducida preserva el concepto de la estructura original contenida en el modelo para la agresividad de Buss y Perry (1992), y puede ser aplicada tanto en hombres como mujeres. Pero al mismo tiempo, mejora la varianza común explicada, y las cargas factoriales permanecen estables en tres diferentes países: Estados Unidos, Canadá e Inglaterra. Con una ventaja adicional determinante: el tiempo de aplicación disminuye sustancialmente (Bryant y Smith, 2001).

Vitoratou *et al.* (2009) revisaron la estructura factorial de la versión de Buss y Perry (1992). Encontraron un inadecuado ajuste para los datos de sus tres muestras griegas: la primera de población no clínica, la segunda de jóvenes que prestaban el servicio militar en la aviación, y la tercera con pacientes drogadictos pertenecientes a un programa con metadona. Sin embargo, al evaluar la estructura del modelo propuesto por Bryant y Smith (2001), los indicadores de ajuste absoluto y relativo fueron comparativamente muy superiores en todas las muestras estudiadas, como se ve en la Tabla 1.

Tabla 1. Indicadores de ajuste del AQ del estudio de Vitoratou *et al.* (2009)

	Muestra	GFI	RMSEA	NNFI	CFI
Buss y Perry (1992) (29 ítems)	1	.71	.096	.57	.61
	2	.86	.067	.76	.78
	3	.77	.074	.73	.75
Bryant y Smith (2001) (12 ítems)	1	.96	.054	.92	.94
	2	.94	.069	.86	.90
	3	.92	.067	.74	.90

Por el contrario, al revisar la consistencia interna obtenida para cada uno de los dos modelos, los coeficientes alfa Cronbach eran mayores para la versión de 29 ítems, como se exhibe en la Tabla 2.

Tabla 2. Consistencia interna del AQ del estudio de de Vitoratou *et al.* (2009)

	Escala	AQ ₂₉	AQ ₁₂
Muestra 1	Física	.82	.77
	Verbal	.55	.55
	Ira	.75	.57
	Hostilidad	.75	.66
	Total	.85	.74
Muestra 2	Física	.82	.70
	Verbal	.50	.56
	Ira	.79	.60
	Hostilidad	.70	.61
	Total	.88	.77
Muestra 3	Física	.84	.76
	Verbal	.51	.48
	Ira	.74	.54
	Hostilidad	.70	.63
	Total	.87	.73

La versión reducida de Bryant y Smith (2001) fue adaptada al castellano por Gallardo-Pujol *et al.* (2006). El constructo de validez de esta estructura factorial fue evaluado, cada ítem cargó en un solo factor y como se esperaba, todos los factores estuvieron correlacionados. La muestra exhibió un buen ajuste a los cuatro factores ($RMSEA=.053$, $CFI=.93$, $TLI=.91$). Y al compararla con la transcultural española de

20 ítems, se observa que la mayoría de los ítems (9 de 12) son estables entre las dos. Los tres que difieren son el ítem 8 (*En ocasiones siento que la vida me ha tratado injustamente*), el ítem 12 (*Parece que siempre son otros los que consiguen las oportunidades*) de la escala *hostilidad*, y el ítem 25 (*Tengo dificultades para controlar mi genio*) de la escala *ira*. (Morales-Vives, Codorniu-Raga y Vigil-Colet, 2005). Los coeficientes de fiabilidad fueron para *agresividad física* de $\alpha=.70$, *agresividad verbal* de $\alpha=.58$, *ira* de $\alpha=.66$, *hostilidad* de $\alpha=.68$ y para el total de $\alpha=.78$ (Gallardo-Pujol *et al.*, 2006). Como se ve, los coeficientes son menores a los publicados en idioma español con todos los ítems por Andreu *et al.* (2002) y García-León *et al.* (2002). Pero según los autores, esto se debe a que el coeficiente alfa de Cronbach se incrementa cuando hay un mayor número de ítems (Gallardo-Pujol *et al.*, 2006). Sin embargo, la consistencia interna es similar a las informadas para las tres muestras de Bryant y Smith (2001).

2.5.4 La versión reducida española del test AQ consistente entre culturas

Vigil-Colet *et al.* (2005) observando que las adaptaciones en distintos idiomas presentaban un inadecuado ajuste en algunos de los ítems, consideraron que éstos podían ser sensibles a las diferencias existentes entre una cultura y otra, situación puesta ya en evidencia en otras publicaciones (Nakano, 2001; Ramírez, Andreu y Fujihara, 2001). Entonces para resolver este problema, intentaron encontrar una estructura factorial que fuera estable para las culturas analizadas, y que

razonablemente estuviera libre de este tipo de sesgos. De tal forma que se obtuviera un cuestionario en el que sólo quedarán aquellos ítems que presentaran las mejores congruencias. Es de aclarar, que para los autores el objetivo principal no era obtener una versión corta, sino una que mantuviera los ítems que tuvieran la característica de ser estables entre las distintas culturas estudiadas. Para la consecución de su objetivo recurrieron a una metodología que se centraba no tanto en el total de la varianza explicada, como se había realizado en las versiones refinadas, sino en la varianza común del modelo entre las muestras. Utilizaron el método Consensus Direct Oblimin con $\gamma=0$ desarrollado por Lorenzo-Seva, Kiers y Ten Berge (2002), que es recomendado cuando no existe un acuerdo óptimo entre las estructuras factoriales encontradas. Dicho método lleva a cabo una rotación oblicua conjunta de las matrices factoriales de las muestras, con el propósito de obtener un compromiso entre simplicidad y acuerdo óptimo entre todas ellas. Esta rotación permite evaluar el índice de congruencia de las distintas soluciones, así como detectar los ítems que presentan problemas, dando información sobre su posible eliminación

Durante el proceso, debido a que el índice de congruencia inicial que se obtuvo fue relativamente bajo (.87), se eliminaron nueve ítems que presentaban baja congruencia entre las muestras. Como es el caso del ítem 4 (*Me suelo implicar en las peleas algo más de lo normal*) y el ítem 7 (*No encuentro ninguna buena razón para pegar a una persona*) de la escala agresividad física; el ítem 3 (*Cuando la gente me molesta, discuto con ellos*) de la escala agresividad verbal; el ítem 4 (*Soy una persona apacible*), el ítem 5 (*Algunos de mis amigos piensan que soy una persona impulsiva*) y el ítem 7 (*Tengo dificultades para controlar mi genio*) de la escala ira; y el ítem 2 (*En*

ocasiones siento que la vida me ha tratado injustamente), el ítem 3 (*Parece que siempre son otros los que consiguen las oportunidades*) y el ítem 6 (*Desconfío de desconocidos demasiado amigables*) de la escala *hostilidad*. Luego, con los ítems restantes se repitió el procedimiento Consensus Direct Oblimin con $\gamma=0$, lográndose esta vez una congruencia total mejorada de .91, considerada como satisfactoria. Asimismo, la congruencia de los factores fue para *agresividad física* de .96, *agresividad verbal* de .89, *ira* de .88 y *hostilidad* de .92. El cuestionario final posee 20 ítems, siete de *agresividad física*, cuatro de *agresividad verbal*, cuatro de *ira* y cinco de *hostilidad* (Vigil-Colet *et al.*, 2005).

Cabe recordar, que esta versión contiene la mayoría los ítems de Bryant y Smith (2001), de tal forma que no están presentes sólo tres de los doce ítems. Y las diferencias entre ellas, debidas a los ítems comunes y no comunes, según Vigil-Colet *et al.* (2005) se deben, primero, a que cada una recurrió a un tipo de análisis factorial diferente, en este caso, el uno confirmatorio y el otro exploratorio. Y segundo, a que la versión española tiene como meta lograr una estabilidad entre muestras de diferentes culturas, mientras que la reducida, pretende explicar la mayor cantidad de varianza posible con el menor número de ítems. Hay que agregar, que varios de los ítems que se eliminaron han sido problemáticos en otros estudios, como es el caso de los invertidos, que por su baja congruencia fueron descartados (Nakano, 2001; Andreu *et al.*, 2002; Vigil-Colet *et al.*, 2005).

En una publicación posterior, Morales-Vives *et al.* (2005) con una nueva muestra de 237 sujetos adultos con niveles educativos muy diversos, analizaron la

consistencia interna y la bondad del ajuste de varios modelos, tanto para la versión completa de 29 ítems de Buss y Perry (1992), como para las reducidas. Los resultados revelaron que la original y las reducidas obtuvieron un aceptable ajuste al modelo de cuatro factores. La adaptación española de Andreu *et al.* (2002) de 29 ítems obtuvo índices de $RMSEA=.055$, $CFI=.88$, $GFI=.84$ y $AGFI=.82$, y una consistencia interna para el total de la escala de $\alpha=.88$, *agresividad física* de $\alpha=.82$, *agresividad verbal* de $\alpha=.77$, *ira* de $\alpha=.67$ y *hostilidad* de $\alpha=.75$. La de Bryant y Smith (2001) indicadores de $RMSEA=.049$, $CFI=.91$, $GFI=.95$ y $AGFI=.92$, y fiabilidades para el total de $\alpha=.81$, *agresividad física* de $\alpha=.81$, *agresividad verbal* de $\alpha=.63$, *ira* de $\alpha=.66$ y *hostilidad* de $\alpha=.68$.

Y en lo que respecta a la de Vigil-Colet *et al.* (2005) los indicadores de ajuste fueron $RMSEA=.056$, $CFI=.90$, $GFI=.89$ y $AGFI=.86$. Con coeficientes alfa Cronbach para el total de $\alpha=.80$, *agresividad física* de $\alpha=.88$, *agresividad verbal* de $\alpha=.71$, *ira* de $\alpha=.68$ y *hostilidad* de $\alpha=.65$. Es de notar que la consistencia de la escala total estuvo muy por encima de la de Bryant y Smith (2001), lo mismo ocurrió con *agresividad verbal* e *ira*, pero con menores diferencias. Y en oposición, *hostilidad* fue la que menor coeficiente obtuvo de las tres que se compararon, por lo que los resultados en esta escala deben ser interpretados con cuidado. Ante estos resultados los autores recomiendan preferir las versiones cortas, por requerir un menor tiempo y esfuerzo en su aplicación y procesamiento. Además, con referencia a la de Vigil-Colet *et al.* (2005), señalan que ésta posee una ventaja adicional, por el procedimiento llevado a cabo en su desarrollo, ya que ofrece una mayor validez transcultural que la de Bryant y Smith (2001), lo que permite llevar a cabo mejores comparaciones cuando se trabaja

con muestras pertenecientes distintos contextos socioculturales (Morales-Vives *et al.*, 2005).

Una característica importante que vale la pena mencionar, es que esta versión se encuentra libre sesgos asociados al sexo, al haberse eliminado los ítems con bajos índices de discriminación relacionados con esta variable, por tanto, permite abordar sin distorsiones las posibles diferencias reales existentes entre hombres y mujeres (Condon, Morales-Vives, Ferrando y Vigil-Colet, 2006; Morales-Vives *et al.*, 2005).

Sobre la base de lo enunciado anteriormente, esta versión libre de sesgos culturales y de sexo, ha sido seleccionada como base para realizar una adaptación en el contexto suramericano, específicamente el colombiano, con preadolescentes y adolescentes, que es uno de los objetivos planteados para la presente tesis.

2.6 Adaptaciones del test AQ en niños y adolescentes

Aunque originalmente diseñado para adultos, sus propiedades han sido evaluadas en edades más tempranas, e incluso se han desarrollado versiones adaptadas exclusivamente a las características de estas edades. No obstante, como se describirá, es común que las propiedades psicométricas no presenten los mismos estándares que en adultos.

A continuación se revisarán los estudios con niños y adolescentes en otros idiomas distintos al castellano. En las versiones largas está la de Morren y Meesters

(2002), que con adolescentes hombres holandeses delincuentes vinculados a un programa de rehabilitación, examinaron la validez de constructo del AQ, a partir de la de la versión para adultos de Meesters *et al.* (1996). Aunque no suministraron datos sobre la estructura factorial, sí reportaron la consistencia interna para *agresividad física* de $\alpha=.75$, *agresividad verbal* de $\alpha=.51$, *ira* de $\alpha=.67$ y *hostilidad* de $\alpha=.69$, y para la escala total de $\alpha=.86$, valores de hecho bastante similar a los obtenidos con adultos. Se observa que la única escala que mostró una consistencia aceptable fue *agresividad física*. *Hostilidad* e *ira* presentaron unas consistencias inferiores, más *agresividad verbal* tuvo un valor pobre.

Fossati *et al.* (2003) han utilizado el análisis confirmatorio, en un estudio cuyo objetivo era medir la replicabilidad de la estructura factorial del AQ. Con una muestra de adolescentes y otra de universitarios, hallaron una estructura invariante, y confirmaron el modelo de cuatro factores. Informaron además una consistencia interna de $\alpha=.81$ para *agresividad física*, $\alpha=.53$ para *agresividad verbal*, $\alpha=.72$ para *ira* y $\alpha=.68$ para *hostilidad*.

En lo relacionado a las versiones refinadas en idiomas diferentes al español se tiene la de Hornsveld, Muris, Kraaimaat y Meesters (2009) que con estudiantes de secundaria informaron un pobre ajuste al modelo con el número original de ítems ($GFI=.72$, $CFI=.65$, $RMSEA=.09$), incluso verificaron los dos factores propuestos por Williams *et al.* (1996), pero hallaron unos indicadores inadecuados ($GFI=.73$, $CFI=.68$, $RMSEA=.09$). En cambio, el modelo de Bryant y Smith (2001) fue el que presentó el mejor ajuste de las tres las soluciones encontradas ($GFI=.93$, $CFI=.94$,

$RMSEA=.06$). A pesar de este ajuste superior, los valores alfa Cronbach fueron contradictorios cuando se contrastó la original con la refinada. Como se puede ver a continuación, la versión original obtuvo un mejor puntaje en el total de la escala ($\alpha=.87$ vs $\alpha=.81$), *agresividad física* ($\alpha=.78$ vs $\alpha=.62$) e *ira* ($\alpha=.64$ vs $\alpha=.51$). Más no ocurrió lo mismo con *agresividad verbal* ($\alpha=.34$ vs $\alpha=.59$) y *hostilidad* ($\alpha=.73$ vs $\alpha=.74$) donde la versión reducida fue superior.

En esta misma línea de las reducidas encontramos en idioma chino a Ang (2007) con dos muestras de adolescentes, quien utilizando la versión de Bryant y Smith (2001) realizó análisis confirmatorios para comprobar cuál modelo era el más adecuado. De los modelos analizados, la versión refinada fue la que mejor ajuste presentó en las muestras. Para la primera muestra, los indicadores fueron de $CFI=.96$, $GFI=.96$, $SRMR=.043$ y $RMSEA=.047$, y para la segunda de $CFI=.98$, $GFI=.97$, $SRMR=.039$ y $RMSEA=.037$. Exhibiendo ajustes comparables a los descritos por Bryant y Smith (2001). Pero la consistencia interna en las dos muestras fue contradictoria y baja en casi todas las escalas, para *agresividad física* ($\alpha=.56$, $\alpha=.72$), *verbal* ($\alpha=.50$, $\alpha=.70$), *ira* ($\alpha=.34$, $\alpha=.55$), *hostilidad* ($\alpha=.38$, $\alpha=.61$) y el total ($\alpha=.84$, $\alpha=.82$).

Ahora bien, en lengua castellana se hallan estudios que replican la estructura factorial original de cuatro factores. López del Pino, Sánchez, Rodríguez y Fernández (2009) con una muestra de españoles aplicó una versión poco conocida que contiene 40 ítems, al realizar el análisis factorial exploratorio de componentes principales y la rotación oblicua, informó de una estructura tetrafactorial. En cuanto a la consistencia

interna, los coeficientes publicados fueron curiosamente bastante altos, con valores de $\alpha=.85$ para *agresividad física*, $\alpha=.85$ para *agresividad verbal*, $\alpha=.64$ para *ira*, $\alpha=.78$ para *hostilidad* y para el total de $\alpha=.91$.

Santisteban y Alvarado (2009) y Santisteban *et al.* (2007) con una muestra compuesta por 1116 preadolescentes ($M=10.0$, $DT=.73$) y 1092 adolescentes ($M=15.3$, $DT=.90$), realizaron una adaptación de la versión de Buss y Perry (1992) para estas edades, que mantenía los mismos ítems y el formato de respuesta original. A través del análisis factorial confirmaron la estructura de cuatro factores. No obstante, en la muestra con el grupo de preadolescentes el ítem 3 (*Me enfado rápidamente, pero se me pasa enseguida*) y ítem invertido 24 (*Nunca encuentro una razón tan grande como para pegarle a una persona*) obtuvieron bajas cargas factoriales. Lo mismo ocurrió con el ítem 23 (*Desconfío de los desconocidos que vienen como muy amiguitos*) tanto en adolescentes como preadolescentes. Los autores interpretaron este desajuste como una consecuencia de las interpretaciones semánticas y la manera en que los jóvenes perciben estos ítems. Hay que tener en cuenta que la baja carga factorial en el ítem invertido reportada aquí, también fue informada por Nakano (2001). Por otra parte con referencia al ítem 23 Harris (1995) y Meesters *et al.* (1996) sugirieron que éste era una medida más de *desconfianza* que de *hostilidad*. A pesar de las anteriores observaciones, consideraron que todos los 29 ítems debían permanecer, pues era conveniente para poder realizar mejores comparaciones, como en caso de los estudios transculturales. El análisis confirmatorio de toda la muestra ofreció un buen ajuste ($GFI=.92$, $AGFI=.91$, $RMSEA=.047$), y puesto que el GFI y el $AGFI$ estuvieron por encima de .90 y $RMSEA$ por debajo de .05, concluyeron que los datos se ajustaban

bien al modelo de Buss y Perry (1992). Del mismo modo, se vio una buena consistencia interna total en preadolescentes (hombres $\alpha=.87$, mujeres $\alpha=.86$) y adolescentes (hombres $\alpha=.88$, mujeres $\alpha=.88$). La fiabilidad total para toda la muestra fue de $\alpha=.87$, para *agresividad física* de $\alpha=.80$, *agresividad verbal* de $\alpha=.73$, *ira* de $\alpha=.65$ y *hostilidad* de $\alpha=.66$.

Por otro lado, existen también en español investigaciones que indican que tanto el modelo original como la versión reducida presentan un buen ajuste al modelo. Reyna *et al.* (2011) en un estudio en Latinoamérica con adolescentes argentinos ($N=371$) entre 12 y 19 años ($M=14.89$; $DT=1.97$) hicieron inicialmente un análisis exploratorio utilizando ejes principales, igual a lo que se realizó en la original, y la estructura factorial que ofreció mejor ajuste fue una bifactorial. El primer factor incluía ítems de *agresividad física*, *agresividad verbal* e *ira*, y el segundo, ítems de *hostilidad*, *agresividad verbal* e *ira*. Los ítems 3 y 24 tuvieron bajas saturaciones y no fueron incluidos en los análisis posteriores. Seguidamente se llevó a cabo el análisis confirmatorio para tres modelos. El primero, bifactorial sobre el modelo exploratorio previo, el segundo, sobre la versión de Buss y Perry (1992) y el tercero sobre el propuesto por Bryant y Smith (2001). Los peores indicadores de ajuste fueron para el modelo bifactorial ($GFI=.92$, $AGFI=.89$, $SRMR=.049$, $CFI=.98$, $RMSEA=.038$). El modelo de 29 ítems obtuvo indicadores mejores ($GFI=.90$, $AGFI=.88$, $SRMR=.051$, $CFI=.97$, $RMSEA=.042$). Más el reducido fue el que ostentó el mayor ajuste ($GFI=.97$, $AGFI=.96$, $SRMR=.038$, $CFI=.99$, $RMSEA=.027$). La exclusión de los dos ítems mencionados está en línea con lo que han considerado varios autores sobre retirar los ítems problemáticos (Harris, 1995; Meesters *et al.*, 1996; Nakano, 2001). Llama la

atención en este estudio, que los dos ítems problemáticos son los mismos reportados con bajas cargas factoriales por Santisteban *et al.* (2007) en su muestra con preadolescentes. En lo referente a la consistencia interna, la versión de 27 ítems presentó valores para total de $\alpha=.87$, *agresividad física* de $\alpha=.77$, *agresividad verbal* de $\alpha=.68$, *ira* de $\alpha=.67$ y *hostilidad* de $\alpha=.71$. Y la versión corta, coeficientes para el total de $\alpha=.82$, *agresividad física* de $\alpha=.70$, *agresividad verbal* de $\alpha=.55$, *ira* de $\alpha=.48$ y *hostilidad* de $\alpha=.54$ (Reyna *et al.*, 2011).

3 La adaptación de test entre culturas

Actualmente en un mundo globalizado caracterizado por un gran intercambio de información, los avances en la ciencia son comunicados cada vez con mayor velocidad. Esta situación ha llevado a incrementar el interés por los estudios transculturales, y la adaptación de test no es la excepción. Cada vez es más frecuente que los instrumentos desarrollados para unos determinados contextos lingüísticos y culturales, sean adaptados en otras poblaciones con características muy distintas a la cultura donde se originó (Balluerka, Gorostiaga, Alonso-Arbiol y Haranburu, 2007). Existen dos razones fundamentales por las cuales se llevan a cabo estas adaptaciones, la primera, es de índole práctica, es decir, adaptar un test es más rápido y económico que desarrollar uno nuevo. Adicionalmente, hay una serie de investigaciones previas que dan soporte a la validez del test que se pretende adaptar. En el caso de Latinoamérica, por su situación económica y por consiguiente, el escaso presupuesto dedicado a la investigación, la gran mayoría de los test que se utilizan son originarios

de Estados Unidos, seguidos por los elaborados en Europa, especialmente España. La segunda razón, es que para realizar estudios que involucran muestras de diferentes culturas, se requieren instrumentos adecuados para poder comparar los resultados. (Ercikan, 2002; Hambleton y Patsula, 1998; Hui y Triandis, 1985; Van de Vijver y Hambleton, 1996).

3.1 Equivalencia

Equivalencia se refiere al nivel de comparación que se puede hacer entre medidas obtenidas en contextos distintos. Involucra elementos de tipo conceptual, cultural, lingüístico y de medida. Para obtener una adecuada equivalencia hay que cumplir una serie de condiciones. Primera, debe haber una equivalencia del constructo entre las culturas, lo que significa que el significado sea compartido entre los grupos. Segunda, el instrumento tiene que tener la misma estructura factorial en los grupos estudiados. Tercera, el test debe poseer la misma unidad de medida en los grupos (Barbero, Vila y Holgado, 2008; Hambleton, 2005; Muñiz y Hambleton, 1996; Van de Vijver y Leung, 2000).

3.2 Sesgos en el proceso de adaptación

El hecho que un test haya demostrado a cabalidad sus propiedades psicométricas en una determinada cultura, no implica que automáticamente pueda ser aplicado a otra. Desarrollar un test con cualidades psicométricas aceptables para otro grupo cultural

requiere mucho más un simple esfuerzo de traducción literal (Muñiz, 1994). Incluso si el idioma es el mismo, como en el caso del español, las características idiomáticas y lingüísticas pueden variar sustancialmente, y aunque se utilicen exactamente las mismas palabras, el significado y la interpretación pueden ser muy distintos. Cuando un test psicológico ha sido desarrollado en una cultura y se aplica en otra, la fiabilidad y la validez no pueden ser asumidas a priori, hay que demostrarlo empíricamente. Y aunque las propiedades psicométricas para cada una de las versiones separadamente sean excelentes, esto no es suficiente para realizar comparaciones. Se precisa que el significado psicológico sea el mismo, es decir, que exista una correspondencia entre los puntajes observados de los sujetos de las culturas estudiadas. Sólo en este caso, puede afirmarse que las variables se encuentran libres de sesgos, y por tanto son equivalentes (Hambleton, 2001, 2005; Van de Vijver y Poortinga, 2005).

Al llevar a cabo el proceso de adaptación, es posible que se presenten sesgos que pueden afectar la validez. Éstos han sido clasificados de acuerdo a sus características en tres tipos: sesgos de constructo, sesgos de método y sesgos de los ítems (Van der Vijver y Hambleton, 1996).

3.2.1 Sesgos de constructo

Ocurren cuando el constructo no es idéntico entre los grupos comparados. Este sesgo tiene su origen en diferentes fuentes. La primera, porque las definiciones de los constructos no son exactamente las mismas para cada una de las culturas, es decir, no existe una concordancia total entre ellos. Segunda, debido a que la apropiación de los

comportamientos esperados asociados a los constructos es diferente entre los grupos culturales. Y tercera, porque los ítems no abarcan adecuadamente las conductas y elementos más relevantes relacionados con el constructo. Si estos sesgos no son tenidos en cuenta, y no se toman las medidas del caso para controlarlos, la adaptación perdería sentido, ya que se estarían evaluando constructos psicológicos distintos en cada una de las culturas (Carretero-Dios y Pérez, 2005; Van de Vijver y Tanzer, 2004).

3.2.2 Sesgos de método

Pueden producirse en tres situaciones. La primera, cuando las muestras no son homogéneas y no admiten realizar comparaciones válidas. La segunda, porque existen diferencias en la forma de aplicación del instrumento. La tercera relacionada con las características del instrumento, causada por las diferencias referentes a la familiaridad con los ítems y al estilo de respuesta (Van de Vijver y Poortinga, 2005; Van de Vijver y Tanzer, 2004).

3.2.3 Sesgos de los ítems

Un ítem se encuentra sesgado cuando individuos con características similares, pero pertenecientes a culturas distintas, no obtienen puntajes similares. Un ítem sin sesgo debe tener el mismo sentido significado y grado de dificultad, para cada variable que se pretende medir. Por lo tanto, las medias de los puntajes deben ser

consistentemente similares entre las culturas. (Holland, y Thayer, 1988; Shepard, Camelli y Averill, 1981).

Dentro de la adaptación de un ítem hay que garantizar que el lenguaje sea el adecuado en sus aspectos idiomáticos y culturales. El vocabulario y tipo de redacción utilizado tiene que ser comparable con respecto al nivel de dificultad, complejidad, estilo, longitud, etc. Igualmente hay que verificar si cada uno de los ítems tiene el mismo significado, y se ha de tener en cuenta que no siempre las palabras y expresiones son las mismas en la otra cultura (Hambleton, 2005; Muñiz y Hambleton, 1996; Van de Vijver y Poortinga, 1997, 2005).

Ampliando lo anterior, los sesgos de los ítems pueden ser debidos a varias fuentes. La primera, causada por una inadecuada traducción, o una traducción literal que no tiene en cuenta la el sentido y el significado de las palabras y expresiones idiomáticas dentro de cada lenguaje. O si es en el mismo idioma, no se consideran los diferentes usos y significados idiomáticos de cada región, ya que la misma palabra o frase puede tener connotaciones muy distintas entre una cultura y otra. La segunda, por la escasa familiaridad o conocimiento del contenido del ítem en una de las culturas (Sireci, 2011; Van der Vijer y Hambleton, 1996; Van de Vijver y Leung, 2000).

Al respecto Hambleton y Zenisky (2011), señalan que en la adaptación de un test, los investigadores deben hacerse cuatro cuestionamientos para evaluar el proceso realizado en cada ítem. El primero, se relaciona con la verificación de si el significado

es el mismo, tanto en la cultura original como en la cultura en que se realiza la adaptación. Ya que es posible que la forma en que se entiende y responde el ítem en cada una de las culturas sea diferente. El segundo, busca verificar si los ítems tienen el mismo nivel de dificultad y familiaridad. El tercero, evalúa si los cambios introducidos en el texto, tales como omisiones, sustituciones y adiciones, pueden afectar la dificultad de los ítems. Y el cuarto, busca identificar si existen diferencias, cuando se utilizan metáforas, modismos, expresiones idiomáticas, etc.

3.3 Directrices para la adaptación de test

Como son numerosos los problemas que se observan en las publicaciones que reportan la adaptación de instrumentos, en lo referente a la idoneidad de los procedimientos seguidos, se han desarrollado diversas guías en los últimos años. No obstante, buscando unificar el avance teórico, las experiencias y pautas propuestas por diferentes autores y asociaciones, la *International Test Commission* ha elaborado unas pautas para la adaptación de test. Esta guía contiene 22 directrices que han sido organizadas en cuatro categorías. La primera, el contexto, que se orienta en la equivalencia del constructo entre las culturas estudiadas. La segunda, construcción y adaptación, centrada en todo el proceso de adaptación, desde la selección de los traductores, hasta los métodos estadísticos de análisis para verificar la equivalencia de los puntajes. La tercera, relacionada con el proceso de aplicación. Y la cuarta, referida a la interpretación de las puntuaciones (Hambleton, 2001, 2005; Muñiz y Hambleton, 1996; Van de Vijver y Hambleton, 1996).

OBJETIVOS

1 Objetivos relacionados con el test BIS-11c

El objetivo general de esta parte de la investigación es adaptar la versión española de la Escala Barratt de Impulsividad de para niños de Cusi *et al.* (2008b) a las características culturales y lingüísticas propias de los niños colombianos.

Este objetivo se encuentra conformado por los objetivos específicos que se enuncian a continuación:

- Contrastar a nivel de análisis factorial exploratorio hasta qué punto la estructura colombiana y la española son congruentes entre sí.
- Comprobar mediante análisis factorial confirmatorio si la solución resultante de la adaptación colombiana replica los tres factores propuestos por Barratt (1985).
- Comprobar hasta qué punto la estructura colombiana y la española comparten las mismas características en términos de invarianza configuracional (*configurational invariance*), invarianza factorial débil (*weak factorial invariance*) e invarianza factorial fuerte (*strong factorial invariance*).
- Determinar la consistencia interna de las escalas *impulsividad motora*, *impulsividad cognitiva* e *impulsividad no planificadora*.

- Determinar la relación existente entre las tres escalas (*impulsividad motora, impulsividad cognitiva e impulsividad no planificadora*), y en qué casos ésta es positiva o negativa.

2 Objetivos relacionados con test AQ

El objetivo general para esta parte de la investigación es adaptar la versión española para adultos del Cuestionario de Agresividad de Buss y Perry, desarrollado por Vigil-Colet *et al.* (2005), a las características culturales y lingüísticas de los preadolescentes y adolescentes colombianos.

El objetivo se encuentra formado por los objetivos específicos que a continuación se describen:

- Contrastar a través del análisis confirmatorio si la solución factorial obtenida de la muestra colombiana replica los cuatro factores propuestos por Buss y Perry (1992).
- Determinar la consistencia interna total y de las escalas *agresividad física, agresividad verbal, ira y hostilidad* para toda la muestra.
- Determinar la consistencia interna de cada una de las escalas en los grupos de edad de 8 y 10 años, 11 y 13 años y 14 y 16 años.

- Comprobar si existen diferencias significativas entre los valores obtenidos por los varones y las mujeres, en las distintas escalas del instrumento.

MÉTODO

A continuación se presentan los diversos pasos y procedimientos metodológicos que se han desarrollado para la elaboración de este estudio. Inicialmente se describe la muestra utilizada, posteriormente se explican las características de las versiones españolas de los test BIS-11c y AQ, para luego relatar con detalle cómo fue realizado el proceso de adaptación cultural que llevó a obtener las respectivas versiones para niños y adolescentes colombianos. Finalmente, después de explicar cómo realizó la aplicación de los instrumentos, se expone el tipo de estadísticos utilizados, en especial los relacionados con el análisis factorial.

1 Participantes

Se ha trabajado con una muestra incidental. Estuvo compuesta por un total de 616 estudiantes de los cuales 306 (49.7%) son varones y 310 (50.3%) mujeres, pertenecientes a tres institutos de la ciudad de Bucaramanga (Colombia). Las edades estuvieron comprendidas entre los 8 y 16 años ($M=13.0$, $DT=2.38$). Los alumnos se encontraban cursando grados entre 4° y 11°. Se eliminaron de las muestras finales los sujetos que no cumplieron adecuadamente los instrumentos. Finalmente la muestra para la adaptación del BIS-11c fue de 538 niños, mientras que la utilizada en el AQ estuvo compuesta por 535 niños.

Teniendo en cuenta que la mayor parte de la población de niños asiste a institutos oficiales, se seleccionaron dos oficiales y uno privado. Fueron elegidos siguiendo también un criterio de distribución geográfica, por lo cual se ha escogido de los institutos oficiales uno que atiende la población que vive en centro-norte de la

ciudad y otro que atiende la del centro-sur, siendo éstos clasificados dentro del grupo de los que más alumnos atienden en la ciudad, alrededor de 1600 niños distribuidos en las jornadas de mañana y tarde.

Los participantes de los institutos oficiales pertenecían al Centro Piloto Simón Bolívar ($N=273$) localizado en la zona norte de la ciudad en la Comuna 3 (San Francisco), que atiende los barrios: Norte bajo, San Rafael, El Cinal, Chapinero, Comuneros, La universidad, Mutualidad, Modelo, San Francisco y Alarcón. Por su parte el Instituto INEM Custodio García Rovira ($N=140$), se encuentra localizado al sur de la ciudad en la Comuna 10 (Provenza), atiende alumnos pertenecientes a los barrios: Diamante II, San Luís, Provenza, El cristal, Fontana y Granjas de Provenza. Por su carácter de institución oficial y su localización, estos institutos reciben en sus aulas alumnos de todos los niveles socioeconómicos, siendo los estratos bajo y medio los que predominan.

De otro lado, el instituto privado seleccionado, el Gimnasio Superior ($N=203$), se encuentra situado en el centro de la ciudad en la Comuna 6 (La Concordia), que atiende los barrios: La Concordia, San Miguel, Candiles, Aeropuerto Gómez Niño, Ricaurte, La Ceiba, La Salle y La Victoria. Por su carácter privado la mayor parte de los estudiantes pertenecen al estrato socioeconómico medio.

2 Instrumentos

2.1 *El test BIS-11c*

La escala utilizada como base para la adaptación colombiana del BIS-11c es la versión desarrollada por Cosi *et al.* (2008b) para niños españoles. El instrumento al igual que el original de Barratt (1985) posee tres factores: *impulsividad motora* (13 ítems) asociada a la tendencia de actuar sin pensar, *impulsividad no planificada* (8 ítems) que implica falta de planificación del futuro, e *impulsividad cognitiva* (5 ítems) relacionada con pensar y tomar decisiones rápidamente. Está compuesto por 26 ítems en escala tipo Likert, posee el mismo formato de respuestas de la escala para adultos, con cuatro opciones, donde un puntaje mayor implica más impulsividad (nunca/casi nunca, algunas veces, a menudo, siempre/casi siempre).

La consistencia interna de las escalas es para *impulsividad motora* de $\alpha=.80$, *impulsividad no planificadora* de $\alpha=.73$ e *impulsividad cognitiva* de $\alpha=.68$.

2.2 *El test AQ*

Para la adaptación colombiana del AQ se ha tomado la versión reducida para adultos de Vigil-Colet *et al.* (2005) que en su estructura factorial mantiene los cuatro factores propuestos por Buss y Perry (1992). Fue desarrollada por estos autores eliminando todos aquellos ítems de la versión original que eran sensibles a las

diferencias culturales y que presentaban una baja congruencia entre las distintas culturas.

El instrumento con 20 ítems posee cuatro escalas a saber: *agresividad física* (7 ítems), *agresividad verbal* (4 ítems), estas dos corresponden al componente instrumental, *ira* (4 ítems), asociada con el componente afectivo, y *hostilidad* (5 ítems), que se relaciona con el componente cognitivo.

La consistencia interna para *agresividad física* es de $\alpha=.88$, *agresividad verbal* de $\alpha=.71$, *ira* de $\alpha=.68$ y *hostilidad* de $\alpha=.65$, y para el total de $\alpha=.80$.

3 Procedimiento para la adaptación de los instrumentos

3.1 Procedimiento para la adaptación del test BIS-11c

En una primera etapa un psicólogo colombiano y dos españoles con experiencia y conocimiento en adaptación de instrumentos, revisaron la totalidad de los ítems con el objeto de verificar si eran apropiados cultural y lingüísticamente para la población colombiana entre 8 y 16 años, buscando en especial que el sentido y significado de cada uno de los ítems permaneciera, de tal forma que la nueva versión midiera los mismos constructos de impulsividad que la española. En este proceso, después de

analizar las distintas opciones posibles, se realizaron ajustes en 21 de los 26 ítems de la versión española de Cosi *et al.* (2008b).

Los cambios tuvieron diversos grados de modificación sobre los ítems originales. En este sentido algunos fueron sometidos sólo a ligeras variaciones, que se consideró ayudaban a una mejor comprensión. El ítem *Pienso con rapidez* se modificó por *Pienso rápidamente*, el ítem *Cambio frecuentemente de amigos* por *Cambio con frecuencia de amigos*, el ítem *Decido las cosas rápidamente* por *Decido rápidamente*, el ítem *Ahorro todo lo que puedo* por *Ahorro lo que más puedo* y el ítem *Soluciono los problemas uno a uno* por *Soluciono los problemas uno por uno*.

Otros ítems tuvieron unos cambios algo mayores. Por ejemplo, el ítem *Hago proyectos para el futuro* fue reemplazado por *Hago planes para el futuro*, el ítem *Me cuesta estar atento* por *Me cuesta trabajo estar atento* y el ítem *Cambio a menudo de ideas* por *Cambio con facilidad mi manera de pensar*.

Y finalmente, unos ítems han tenido una transformación drástica, en la que se han debido cambiar gran parte de las palabras, para construir una oración diferente que preserve el sentido y significado de la versión española. El ítem *Me distraigo con facilidad cuando tengo un problema complicado* fue sustituido por *Cuando estoy haciendo algo que requiere concentración, me distraigo con facilidad*, el ítem *Pierdo los nervios con facilidad* por *Me descontrolo con facilidad*, el ítem *Estoy inquieto en el cine o en clase* por *Me cuesta trabajo quedarme quieto en clase*, el ítem *Cambio a*

menudo de ideas por *Cambio con facilidad mi manera de pensar* y el ítem *Actúo impulsivamente (sin pensar)* por *Actúo sin pensar*.

El siguiente paso consistió en solicitarles a ocho licenciados en educación dedicados a la docencia en distintos institutos de la región, que revisaran y evaluaran detenidamente cada uno de los ítems propuestos en cuanto a la facilidad de comprensión para estas edades, y si era el caso, propusieran alternativas de redacción. Sus observaciones condujeron a algunos cambios adicionales. Por ejemplo, el ítem *Planifico mi tiempo libre* fue sustituido por *Organizo mi tiempo libre*, el ítem *Me descontrolo con facilidad* por *Me desespero con facilidad*, el ítem *Me gusta pensar detenidamente las cosas* por *Me gusta pensar bien las cosas* y el ítem *Planifico mis actividades* por *Organizo mis actividades*.

Una vez realizados todos los cambios y ajustes mencionados, se obtuvo la versión final del BIS-11c para ser administrada a la muestra de niños colombianos.

3.2 Procedimiento para la adaptación del test AQ

De manera similar al procedimiento anterior, en una primera fase un psicólogo colombiano y dos españoles con experiencia y conocimiento en adaptación de instrumentos revisaron los 20 ítems del cuestionario español que originalmente fue desarrollado para adultos. Teniendo cuidado que su adecuación cultural y lingüística fueran realmente apropiados para las edades de la muestra de niños, de tal forma que se salvaguardara el mismo significado y sentido en las dos versiones. El resultado fue

el cambio de 16 de los 20 ítems de la versión española para adultos de Morales-Vives *et al.* (2005).

Los ítems fueron sometidos a diversos niveles de modificación. En primer término se encuentran aquellos que tuvieron variaciones leves, que se orientaron a cambiar sólo algunas palabras. El ítem *A menudo no estoy de acuerdo con la gente* fue modificado por *Con frecuencia no estoy de acuerdo con la gente*, el ítem *Algunas ocasiones siento que la gente se está riendo de mí a mis espaldas*, por *A veces siento que la gente se ríe de mí a mis espaldas* y el ítem *Cuando no estoy de acuerdo con mis amigos, discuto abiertamente con ellos* por *Cuando no estoy de acuerdo con mis amigos, discuto con ellos*.

Otros ítems tuvieron cambios un poco más grandes. El ítem *Si tengo que recurrir a la violencia para proteger mis derechos, lo hago* fue reemplazado por *Si tengo que pelear para defender mis derechos, lo hago* y el ítem *He amenazado a gente que conozco* por *He intimidado a personas que conozco*.

Y por último, otros fueron sometidos a modificaciones mucho más drásticas. El ítem *Algunas veces pierdo los estribos sin razón* fue sustituido por *Algunas veces se me “salta la piedra” sin razón*, el ítem *He llegado a estar tan furioso que rompía cosas* por *Cuando me han “sacado la piedra” he dañado cosas* y el ítem *Cuando estoy frustrado, muestro el enfado que tengo* por *Cuando tengo rabia, no la disimulo*.

Al presentar la versión preliminar a los ocho licenciados en educación, éstos propusieron realizar cambios a algunos ítems. El ítem *Si me molestan mucho, puedo llegar a golpear a otra persona*, fue reemplazado por *Si me molestan mucho, puedo llegar a pegarle a otra persona*, el ítem *Si alguien me golpea, le respondo golpeándole también* por *Si me pegan, yo devuelvo el golpe* y el ítem *Hay personas que me provocan tanto que terminamos pegándonos* por *Hay personas que me molestan tanto que terminamos pegándonos*.

Finalmente, después de todo este proceso, se obtuvo la versión final del AQ para ser aplicada a la muestra de niños colombianos.

4 Procedimiento para la administración de los dos instrumentos

Inicialmente se llevaron a cabo una serie de reuniones con los rectores, coordinadores de jornada, directores de curso y psico-orientadores de los tres centros educativos seleccionados (Gimnasio Superior, Instituto Nacional de Educación Media Custodio García Rovira y Centro Piloto Simón Bolívar), con el fin de informar sobre los objetivos, alcances del estudio y procedimiento que se llevaría a cabo, e igualmente obtener los permisos respectivos para recolectar los datos.

El día señalado, en horario escolar, antes de la entrega de los cuestionarios, se explicó brevemente a cada grupo de alumnos en qué consistía el estudio, solicitando la

colaboración voluntaria. Los instrumentos fueron administrados por un psicólogo y una psicóloga, en compañía del profesor de turno, en los salones de clase donde tenían las actividades académicas los alumnos, y dentro del horario escolar habitual. Los grupos tenían un promedio de 30 alumnos. Las instrucciones, condiciones de aplicación y manejo de la información, garantizaban la confidencialidad y el anonimato de todos los datos suministrados. El tiempo para resolver los dos cuestionarios osciló entre 12 y 15 minutos.

5 Procedimiento para el análisis de datos

5.1 Procedimiento para el análisis de datos del test BIS-11c

En un comienzo se calcularon los estadísticos descriptivos, a continuación se procedió a aplicar el análisis factorial exploratorio de los datos de la muestra colombiana. Se calculó la matriz policórica de correlaciones, realizándose posteriormente la extracción de factores mediante mínimos cuadrados no ponderados, identificándose una estructura trifactorial. Luego, la estructura factorial colombiana fue rotada, junto con la obtenida previamente en la muestra española (Cosi, 2008; Cosi *et al.*, 2008b), utilizando el método Consensus Direct Oblimin (Lorenzo-Seva *et al.*, 2002). Las cargas de las matrices fueron oblicuamente rotadas para satisfacer los criterios de simplicidad y acuerdo.

Para evaluar el grado de similaridad entre las dos estructuras factoriales, la colombiana y la española, se ha calculado el índice de congruencia Tucker (1951) entre los 26 ítems, los tres factores y las dos estructuras. Igualmente, el grado de simplicidad de cada uno de las soluciones factoriales fue medido por medio del índice de simplicidad factorial (*loading simplicity index*), el cual señala hasta qué punto los ítems tienen cargas altas en un factor y nulas o bajas en otros, de tal forma que la solución pueda ser fácilmente interpretable (Lorenzo-Seva, 2003).

A nivel de análisis factorial confirmatorio, se calcularon los índices de bondad de ajuste del modelo, que para este caso fueron: χ^2 , *RMSEA* y *CFI*. Pero como también se quería evaluar la invarianza factorial entre la solución colombiana y la española, se utilizó ESEM (*exploratory structural equation modelling*). Se calcularon los índices de ajuste para la invarianza configuracional (*configurational invariance*), y para dos modelos más restrictivos como son la invarianza factorial débil (*weak factorial invariance*) y la invarianza factorial fuerte (*strong factorial invariance*) (Asparouhov y Muthén, 2009). Los puntos de corte para los diferentes indicadores fueron los propuestos para estos casos (Bentler, 1990; Browne y Cudeck, 1993; Hu y Bentler, 1999).

En lo referente a la consistencia interna, los coeficientes tanto de la escala total como de las tres subescalas, fueron obtenidos utilizando para tal fin los procedimientos y puntos de corte propuestos por Cronbach (1951).

Los programas estadísticos utilizados han sido el Mplus 5.1 (Muthen y Muthen, 2008), Factor (Lorenzo-Seva y Ferrando, 2006) y SPSS 17.0.

5.2 Procedimiento para el análisis de datos del test AQ

Inicialmente han sido calculados los estadísticos descriptivos. Seguidamente, teniendo en cuenta la evidencia empírica anterior tanto en niños como adultos, en los que se ha repetido la estructura de cuatro factores con las escalas de *agresividad física*, *agresividad verbal*, *ira* y *hostilidad*, se ha decidido que un análisis factorial confirmatorio es mucho más pertinente que uno exploratorio para este caso. En consecuencia, se ha tomado como referencia para realizar la confirmación de la solución factorial colombiana, la estructura de la versión de 20 ítems para adultos de Vigil-Colet, *et al.* (2005) que replica los cuatro factores originalmente encontrados por Buss y Perry (1992). La evaluación de modelo se ha hecho con el indicador de ajuste absoluto *RMSEA* y los indicadores de ajuste relativo *CFI* y *NFI*, teniendo como referencia las orientaciones dadas para estos casos por Hu y Bentler (1999).

Mediante un análisis de la varianza (ANOVA) se verificó si existían diferencias significativas entre los grupos de sexo y edad.

El cálculo de la consistencia interna para todas las escalas y el total en los tres grupos de edad fue realizado utilizando el coeficiente de consistencia interna propuesto por Cronbach (1951).

Los programas utilizados fueron para este caso fueron el Mplus 5.1 (Muthen y Muthen, 2008) y SPSS 17.0.

RESULTADOS

1 Resultados referentes al test BIS-11c

Los resultados obtenidos con el test BIS-11c se han publicado en: *Chahin, N., Cosi, S., Lorenzo-Seva, U., y Vigil-Colet, A. (2010). Stability of the factorial structure of Barrat's Impulsivity Scales for children across cultures: a comparison between Spain and Colombia. Psicothema, 22, 983-989.*

A continuación se presenta un resumen de los resultados más relevantes, y posteriormente se imprime el artículo tal y como ha sido publicado.

- El proceso de adaptación llevó a que se realizaran una serie de cambios en 21 de los 26 ítems del instrumento. Para el logro de este objetivo se tuvo en cuenta la opinión de expertos psicólogos y de los profesores, para obtener finalmente una versión que se adaptara a las características culturales y lingüísticas de los niños colombianos.
- En lo referente al análisis factorial exploratorio con el método Consensus Direct Oblimin, la congruencia total entre la versión colombiana y la española fue de .94, y la de los factores fue de .97 para *impulsividad motora*, .90 para *impulsividad cognitiva* y .95 para *impulsividad no planificadora*. Los valores de las congruencias factoriales sugieren que las dos soluciones pueden ser considerados como equivalentes.

- La estructura factorial del BIS-11c compuesta por *impulsividad motora*, *cognitiva* y *no planificadora*, hallada en la versión para niños española de Cosi *et al.* (2008b), y que igualmente refleja el modelo teórico planteado originalmente por Barratt (1985), ha sido confirmada en la adaptación colombiana con los siguientes indicadores de ajuste al modelo: $\chi^2/df=2.01$, $RMSEA=.045$ (intervalo de confianza entre .039 y .051) y $CFI=.86$.
- En lo que respecta a la comparación multigrupo realizada con ESEM (*exploratory structural equation modelling*), entre la adaptación colombiana y la española, la invarianza configuracional (*configurational invariance*) ha mostrado una condición invariante en el patrón de la matriz de los tres factores. Los indicadores de ajuste fueron los siguientes: $\chi^2/df=2.05$, $RMSEA=.045$ y $CFI=.88$. Lo que permite aseverar que los tres factores en las dos versiones del BIS-11c son los equiparables.
- En el método más restrictivo, la invarianza factorial débil (*weak factorial invariance*), que indica que la diferencia entre las covarianzas de las variables se deben a factores comunes y no a otra fuentes de asociación, exhibió indicadores que sólo permiten aseverar bajo ciertas interpretaciones que este supuesto se cumple, como es el caso de la comparación entre la diferencia entre invarianza factorial débil y la invarianza configuracional ($CFI=.85$, $RMSEA=.049$).
- La invarianza factorial fuerte (*strong factorial invariance*) no ofreció buenos indicadores ($CFI=.73$, $RMSEA=.061$). Por lo que aquí no puede afirmarse de ninguna

forma que la diferencias entre los puntajes de las dos versiones son atribuibles a diferencias reales del constructo.

- Los coeficientes de fiabilidad, aunque inferiores a los de la versión española, han sido aceptables para *impulsividad motora* con $\alpha=.74$ (intervalo de confianza entre .70 y .77) y para *impulsividad no planificadora* con $\alpha=.72$ (intervalo de confianza entre .68 y .75). Pero *impulsividad cognitiva* ha reportado valores inaceptables con $\alpha=.59$ (intervalo de confianza entre .55 y .62).
- La correlación informada entre los factores, de la misma forma que la española, ha mostrado que *impulsividad cognitiva* correlaciona negativamente con *impulsividad motora* ($r=-.15$) y con *impulsividad no planificadora* ($r=-.24$). Mientras que *impulsividad motora* correlaciona positivamente con *impulsividad no planificadora* ($r=.35$).

Stability of the factor structure of Barratt's Impulsivity Scales for children across cultures: A comparison of Spain and Colombia

Nicolás Chahín, Sandra Cosi, Urbano Lorenzo-Seva and Andreu Vigil-Colet
Universidad Rovira i Virgili

Despite the great importance of impulsivity in many childhood and adolescence disorders, there are few self-reports on child impulsivity. Recently, a modified version of Barratt's BIS-11 questionnaire adapted for children has been developed, which may be useful in assessing this personality dimension. The present study reports an adaptation of this questionnaire in a different culture (Colombia) and assesses the degree of convergence between the factor structures of both adaptations using consensus oblimin rotation. The results indicate not only that the factor structure of the test remains stable across both adaptations, and that two of the three scales in the Colombian version show acceptable reliabilities, but also that cultural and linguistic issues are important in test adaptation even when the same language is used.

Estabilidad de la estructura factorial de la escala de impulsividad para niños de Barratt a través de culturas: una comparación entre España y Colombia. Apesar de la gran importancia de la impulsividad en múltiples patologías de la infancia y la juventud, existen pocos cuestionarios de impulsividad para niños. Recientemente se ha desarrollado una versión modificada del cuestionario BIS-11 de Barratt adaptado para niños que puede ser útil en la evaluación de esta dimensión de personalidad. Este estudio presenta una adaptación de este cuestionario a una cultura distinta (colombiana) y evalúa el grado de convergencia entre las soluciones factoriales de ambas versiones del test utilizando el método de rotación consensos oblimin. Los resultados indican una elevada estabilidad de la solución factorial a través de las versiones y una fiabilidad satisfactoria para dos de las tres escalas en la adaptación colombiana. Por otra parte se pone de manifiesto la importancia de los aspectos culturales y lingüísticos en la adaptación de un test incluso cuando se trata de la misma lengua.

In recent years there has been an increase in the amount of research on impulsivity. One of the main reasons for this is that impulsivity plays a prominent role in understanding and diagnosing various forms of psychopathology, specially those linked to the lack of impulse control, such as aggression, substance abuse, etc. (Barratt & Slaughter, 1998; Whiteside & Lynam, 2001; McMullan, Blair, & Egan, 2002; Scarpa & Raine, 2002; Vigil-Colet, Morales-Vives, & Tous, 2008). The importance of impulsivity during childhood and adolescence has been established and related to a wide variety of externalising and internalising pathologies, such as hyperactivity, aggression, learning problems, anxiety disorders, depression etc. (Fink & McCown, 1993; Zaparniuk & Taylor, 1997; American Psychological Association, 2000; Willcutt & Pennington, 2000; Summerfeldt, Hood, Antony, Richter, & Swinson, 2004; Cataldo, Nobile, Lorusso, Battaglia, & Molteni, 2005; Jensen, Youngstrom, Steiner, Findling, Meyer et al., 2007).

Despite the great importance of impulsivity in children, there is a lack of self-report measures for this factor. Impulsivity in

children is often measured with rating scales completed by other individuals such as parents or teachers or with behavioural tasks. Self-reports are not as frequently used with children as with adults because it is assumed that they are less accurate at assessing their own behaviours (Fink & MacCown, 1993). Nevertheless, there is evidence to suggest that children between 8 and 12 years old are better informants than parents (Rapee, Barrett, Dadds, & Evans, 1994; Muris, Merckelbach, Van Brakel, & Mayer, 1999; Cosi, Canals, Hernández-Martínez, & Vigil-Colet, 2010). Furthermore, self-report and behavioural tasks seem to measure different components of impulsive behaviour, which indicates that both kind of measures need to be used when assessing impulsivity (Reynolds, Ortengren, Richards, & de Wit, 2005).

Few impulsivity self-report scales have been specifically designed for children. Two that have are Eysenck's I6 impulsivity scale and the children's adaptation of Dickman's impulsivity questionnaire (DII-c) (Eysenck & Eysenck, 1980; Eysenck, Easting, & Parson, 1984; Brunas-Walgstaff, Tilley, Verity, Ford, & Thompson, 1997). The I6 questionnaire was developed by Eysenck and Eysenck (1980) to measure two specific dimensions related to impulsivity (impulsiveness and venturesomeness) which have shown good internal consistencies both for the original English version, and the German and Spanish adaptations although the latter version was tested only in an adolescent sample (Eysenck & Eysenck, 1980; Silva, Martorell, & Clemente, 1987; Stadlet & Janke, 2003).

Fecha recepción: 2-9-09 • Fecha aceptación: 5-12-09
Correspondencia: Andreu Vigil-Colet
Centre de Recerca en Avaluació i Mesura de la Conducta
Universitat Rovira i Virgili
43007 Tarragona (Spain)
e-mail: andreu.vigil@urv.cat

The DII-c scales tried to replicate the dimensions of dysfunctional impulsivity (DI) and functional impulsivity (FI) proposed by Dickman (1990). This inventory, however, showed poor internal consistency with children: in particular, FI was below the commonly accepted standards for a test to be considered sufficiently reliable (Brunas-Walgstaff et al., 1997; Cosi, Morales-Vives, Canals, Lorenzo-Seva, & Vigil-Colet, 2008).

Cosi, Vigil-Colet, Canals, & Lorenzo-Seva (2008) proposed a third approach for assessing impulsivity in children using self-reports: adapting Barrat's impulsivity scales, one of the most widely used self-reports for adults in impulsivity assessment (Patton, Stanford, & Barratt, 1995). Fossati, Barratt, Acquarini, & Di Ceglie (2002) developed an adolescent's version of this scale known as BIS 11-a, consisting of short and very simple items that may also be easily understood by children. Cosi et al. (2008) adapted it into Spanish and found a three-factor structure for the Barratt impulsivity scale for children (BIS-c). This structure comprised the factors of motor (Im), lack of planning (Inp) and cognitive (Ic) impulsivity. Two of the scales showed good or sufficient reliabilities but Ic was below acceptable cut-off points, and showed a reliability of $\alpha = 0.60$. Taking this into account, Cosi, Canals, & Vigil-Colet (2008) developed an improved version of the Ic scale by adding three new items, which improved the reliability ($\alpha = 0.70$). They also found significant relationships between the BIS 11-c scales and measures of aggression and scholastic performance, which have been often associated with impulsivity. Some evidence of the validity of BIS 11-c was therefore given.

The aim of the present study is twofold. The first objective is to adapt the BIS 11-c to a culture other than the Spanish one and to assess its factorial structure. The second objective is to compare the factor structure obtained from the Spanish sample with the factor structure obtained from the Colombian sample. Adapting impulsivity tests so that they can be applied in places such as Colombia acquires particular importance if it is taken into account that the levels of aggression in these countries are high and that impulsivity has an important role in their origin.

The second objective addresses the question of that the extent to which both instruments measure the same constructs in exactly the same way across cultures. As Byrne (Byrne & Watkins, 2003; Byrne, 2008) has pointed out, this is a key issue in cross-cultural adaptation because it refers to the extent that a group of nested equalities between both instruments are accomplished. These equalities refer to such key aspects such as whether both adaptations have equal factor structures in the sense that the number of factors and pattern of indicator-factor loadings are identical across groups (configurational invariance), the extent to which relationships between the scores and latent variables are equivalent (equality of factor loadings, also called weak factorial invariance) for both cultures and the strong factorial invariance in the case of equal indicator intercepts (Meredith, 1993).

One way of testing whether they measure the same constructs is to use the means and covariance structures (MACS) approach proposed by Byrne (Byrne & Stewart, 2006; Byrne, 2008). This approach proposes a set of steps that first establishes the separate model fit for each group, then tests a configurational invariance model as a baseline model and finally tests different nested and more restricted models. Nevertheless, several authors have pointed out that conventional Confirmatory factor analysis (CFA) goodness of fit criteria are too restrictive when applied to most tests, especially in personality research (Ferrando & Lorenzo-Seva, 2000; Asparohov & Muthén, 2009).

One solution to this problem is to use the Exploratory structural equation modelling approach (ESEM) proposed by Asparohov & Muthén (2009). This procedure uses a rotated exploratory factor analysis (EFA) measurement model that is applied to a structural equation model, yielding the goodness of fit indexes usually obtained in SEM, and allowing a step approach to multigroup analysis such as the one proposed by Byrne (2008). Nevertheless this method has an important limitation, because the cut-off values for goodness of fit indexes in ESEM have not yet been established because ESEM studies involve a greater number of estimated parameters than the CFA approach. (Marsh et al., 2009). Furthermore, as Asparohov & Muthén (2009) pointed out ESEM is only able to estimate the measurement invariance of factors conjointly and not the specific invariance for one factor.

A second approach to the problem is based on EFA and involves the simultaneous rotation of the loading matrices obtained in two samples to show a mixture of simplicity and optimal agreement between them. The method used is Consensus Direct Oblimin with $\gamma = 0$ (Lorenzo-Seva, Kiers, & Ten Berge, 2002). In this rotation method, the loading matrices are obliquely rotated together to satisfy two criteria: simplicity and agreement. This method will allow us to assess the degree of agreement between the different factorial solutions for each factor and determine if there is any lack of agreement, which items are responsible for it. It should be mentioned that this analytical approach has been successfully used to compare the factorial structures obtained in personality tests across different cultures (i.e., Vigil-Colet, Lorenzo-Seva, Codorniu-Raga, & Morales, 2005), and similar methods have also been applied in cross-cultural adaptation using other rotational procedures such as Procustes rotation (i.e., Balluerka Gorostiaga, Alonso-Arbiol, & Aramburu, 2007). Nevertheless, being an exploratory approach, it does not provide a goodness of fit index and configurational invariance, weak factorial invariance etc. cannot be tested using nested models.

Taking into account the advantages and limitations of both approaches we will apply both in order to complement the information given by the consensus oblimin method (COM) on the overall (scales) and specific (items) congruences with the goodness of fit indexes and different nested models of the MACS approach using ESEM.

Method

Participants

The participants were 616 children (306 boys and 310 girls) aged between 8 and 16 years with a mean of 13 years ($SD = 2.38$) from one private and two state schools in Bucaramanga (Colombia). They were from medium-low and low social classes, respectively. The cases with missing data ($n = 78$) were removed from the analysis. Bucaramanga is a town of 716,000 in habitants and the capital of the Santander region.

The sample used for comparison (Cosi et al., 2008) consisted of 413 children (186 boys and 227 girls) aged between 9 and 13 years with a mean of 11 years ($SD = .92$). The children came from thirteen schools in Reus (Spain), which were randomly chosen from the state schools and state-subsidized private schools in the town. Reus is a medium-sized town of 100,000 inhabitants.

Procedure

The head teachers of each school were informed about the nature of the research before they authorized the test. Then, the instrument was applied to all the classrooms of the centers to volunteer students guaranteeing anonymity. The test was administered by a psychologist in a collective way to groups around 30 individuals.

Instruments

Barrat Impulsivity Scales-11 for children (Cosi et al. 2008): The BIS-11c consisted of 30 items with a 4-point response format (Never/Almost Never, Sometimes, Often, Always / Almost always). Answers were scored with 0, 1, 2 and 3, respectively. The questionnaire measures motor (Im), non planning (Imp) and cognitive (Icog) impulsivity, and shows internal consistencies of 0.80, 0.73, and 0.68, respectively.

Although test adaptation usually involves translating the test from one language to another, cultural and language differences

that affect test scores are not only a translation (Hambleton, 2005). As a consequence, although Spain and Colombia share the same language, various cultural differences and dialectal variants made it necessary to analyze the items of the original Spanish scale to see if they were appropriate for administration in Colombia. With this purpose, one Colombian and two Spanish psychologists with experience in test adaptations analyzed the degree to which the BIS 11-c items were culturally and linguistically suitable for administration in a Colombian sample. After various modifications in the wording of 21 of the 26 items— for instance we changed «perder los nervios» for «me desespero con facilidad» because the first expression is not commonly used in Colombia—we sent the test to eight Colombian teachers who gave their opinion on whether the items would be understood by Colombian children and, if not, proposed an alternative version of the items. This second phase involved only slight changes such as the use of «organize» instead of «plan». Table 1 shows the Spanish version of BIS 11-c and the alternative items for the Colombian version.

Table 1
Items of the Spanish version of BIS 11-c. Between brackets alternative items for the Colombian adaptation. Mean, Standard deviation (s.d.), item-total correlations (r_{jt}) and Cronbach's alpha if the element is removed ($\alpha-j$)

Item	Media	d.t.	r_{jt}	$\alpha-j$
1. Planifico las cosas que hago (organizo las cosas que hago)	1.52	.89	.41	.69
2. Hago las cosas sin pensarlas	.85	.79	.34	.72
3. Decido las cosas rápidamente (decido rápidamente)	1.42	.92	.22	.47
4. Cuando mis amigos me preguntan algo, puedo responder rápidamente	1.69	.98	.29	.43
5. Me cuesta estar atento (me cuesta trabajo estar atento)	.88	.87	.37	.71
6. Pienso rápidamente (pienso con rapidez)	1.63	.92	.37	.38
7. Planifico mi tiempo libre (organizo mi tiempo libre)	1.57	1.11	.44	.69
8. Pierdo los nervios con facilidad (me desespero con facilidad)	1.20	1.00	.33	.72
9. Me concentro rápidamente	1.70	.98	.26	.45
10. Ahorro todo lo que puedo (ahorro lo que más puedo)	1.68	1.09	.26	.72
11. Me gusta pensar detenidamente las cosas (me gusta pensar bien las cosas)	2.19	.91	.52	.67
12. Hago proyectos para el futuro (hago planes para el futuro)	2.13	.99	.29	.70
13. Digo las cosas sin pensar (digo cosas sin pensar)	.99	.83	.45	.71
14. Soy de los primeros en levantar la mano en clase cuando el profesor hace una pregunta	.97	.87	.21	.48
15. Cambio a menudo de ideas (cambio con facilidad mi manera de pensar)	1.21	.97	.23	.73
16. Actúo impulsivamente [sin pensar] (actúo sin pensar)	.72	.80	.44	.71
17. Me distraigo con facilidad cuando tengo un problema complicado (cuando estoy haciendo algo que requiere concentración, me distraigo con facilidad)	1.00	.88	.41	.71
18. Me dejo llevar por mis impulsos	.96	.88	.37	.71
19. Me gusta pensar las cosas	2.19	.90	.54	.66
20. Cambio frecuentemente de amigos (cambio con frecuencia de amigos)	.55	.88	.15	.74
21. Compró las cosas sin pensar (compro cosas sin pensar)	.58	.84	.40	.71
22. Soluciono los problemas uno a uno (soluciono los problemas uno por uno)	1.96	1.00	.39	.70
23. Gasto más de lo que puedo (gasto más de lo que tengo)	.77	.95	.38	.71
24. Cuando pienso en algo me distraigo fácilmente (cuando estoy pensando en algo me distraigo con facilidad)	.97	.87	.43	.71
25. Estoy inquieto en el cine o en clase (me cuesta trabajo quedarme quieto en el cine o en clase)	.84	.93	.34	.72
26. Planifico mis actividades (organizo mis actividades)	1.93	.97	.55	.66

Data analysis

Consensus Oblimin Method: The data obtained in the Colombian population was factor analyzed as follows: first the polychoric correlation matrix was computed, and then Unweighted Least Squares was performed so that three factors were retained. In order to apply the rotation method described above we used two factor structures: one reported by Cosi, Canals, & Vigil-Colet (2008) obtained in a sample of 456 children, and another obtained in our own sample (i.e., the Colombian sample).

To assess the degree of similarity among factor solutions we computed the averaged Tucker's congruence index (Tucker, 1951) for items, factors and overall factor solution (for the details of the computing see, for example, Chan, Ho, Leung, Chan, & Yung, 1999). A threshold of .85 was used for assess congruence (Lorenzo-Seva & ten Berge, 2006). The degree of factor simplicity of each one of the factor solutions was described by the Loading Simplicity (LS) index (Lorenzo-Seva, 2003).

The steps of the analysis were as follows: first, all the factor solutions were simultaneously rotated by Consensus Direct Oblimin; and then the averaged congruence values for items among samples were computed, and the need to eliminate items was assessed.

Finally, the internal consistency reliabilities of the factor scales were assessed using Cronbach's index.

Factorial Invariance using ESEM: As we have described above this method involves to test the multigroup equivalence of a test using the ESEM approach developed by Asparohov & Muthén (2009). To this end we tested the 3-factor model for each group separately. We then tested the configural invariance that was used as the baseline model for other nested models which tested weak and strong factorial invariance.

We analysed the data using SPSS 17.0, Mplus 5.1 and FACTOR (Lorenzo-Seva & Ferrando, 2006). We used FACTOR for EFA as well as SPSS because it enabled us to use polychoric correlation matrices and gave complementary analyses which are not provided by SPSS.

Results

Tables 1 and 2 show descriptive statistics for the BIS 11-c items and scales. As can be seen, the Colombian sample showed higher values than the Spanish sample for all scales although the differences were significant only for Im ($d=0.22$) and Inp ($d=0.43$). Finally, all the scales, with the exception of Ico ($F_{(1,1120)}$

Scale	Sample	Mean	Std. dev.
Cognitive	Colombia	7.41	2.73
	Spain	7.14	3.09
Non planning	Colombia	15.21	4.57
	Spain	13.29	4.50
Motor	Colombia	12.95	5.42
	Spain	11.65	5.51

p<0.01

= 6.54; $p<0.05$), showed the same variance in the Spanish and Colombian samples. We found that neither sex nor age had any effect on BIS 11-c scales.

COM showed that the overall factor congruence value after the rotation was .94. If we follow the guidelines by Lorenzo-Seva and ten Berge (2006), this congruence value suggests a fair similarity among factor structures. The factor congruence for Ico was .90 (i.e., a fair similarity), whereas the congruence values between the Inp and Im factors were .95 and .97, respectively. These values suggest that these factors can be considered equivalent. Table 3 shows the factor structures in both samples after Consensus Oblimin rotation, and the congruence between items. Most of the items showed a congruence larger than .95, and six items (4, 5, 8, 21, 22 and 23) showed congruence values between .85 and .95 (i.e., a fair similarity). Finally, item 20 showed a congruence of .79, which suggests that it should be deleted. However, the inspection

Items	Spanish sample			Colombian sample			Congruences
	Cognitive	Non-Planning	Motor	Cognitive	Non-Planning	Motor	
1	.07	-.53	.04	.11	-.46	-.05	.98
2	.03	.06	.61	.05	.15	.44	.97
3	.42	.07	.15	.51	.02	.24	.99
4	.34	-.10	-.02	.41	.06	-.12	.89
5	-.28	-.04	.54	-.04	.07	.53	.90
6	.58	-.05	.03	.63	.00	-.03	.99
7	.06	-.51	.11	-.06	-.57	.09	.97
8	-.09	-.21	.45	.06	-.09	.49	.93
9	.43	-.22	-.24	.41	-.15	-.24	.99
10	-.09	-.41	-.10	.01	-.27	-.12	.96
11	-.11	-.55	-.04	-.05	-.61	-.08	.99
12	.17	-.42	.24	.11	-.32	.09	.98
13	.13	.16	.66	-.04	.09	.57	.97
14	.41	-.16	-.17	.34	-.20	-.08	.97
15	-.03	-.18	.47	-.04	-.17	.34	.99
16	.05	.13	.68	.01	.09	.61	1.00
17	-.15	-.11	.53	-.11	-.03	.56	.99
18	.10	.05	.45	.04	.13	.49	.98
19	-.02	-.60	-.22	-.04	-.63	-.18	1.00
20	.03	.02	.30	.19	-.27	.43	.79
21	.15	.06	.48	.03	-.07	.59	.94
22	.07	-.49	-.17	.09	-.51	.01	.94
23	.17	.04	.42	-.01	-.05	.57	.91
24	-.16	-.06	.44	-.20	-.04	.53	1.00
25	-.09	-.10	.52	.00	.03	.53	.96
26	.10	-.63	.09	.03	-.69	.03	.99

of the loading values in the Colombian sample shows that the salient loading was in the same factor as in the Spanish sample so we finally decided to keep it. However, if this problem is replicated in other studies, it could finally be deleted. We concluded that both versions of the test were congruent among samples.

Table 4 shows the consensus inter-factor correlation values for the reduced test. As we can see *Im* and *Inp* are positively related while *Ic* shows a negative pattern of relationships with them. This pattern of relationships is not surprising because impulsive individuals often show inhibition deficits and do not predict consequences (both inhibition and prediction are related to *Im* and *Inp*). On the other hand, *Ic* is more related to quick decisions when this strategy is appropriate than to non reflexive responses.

To assess the level of factor simplicity, we computed the LS index. The values were .34 and .46 for the Spanish and the Colombian samples, respectively. Actually, the index showed that the simplest factor solution was the one obtained in the Colombian sample.

Table 5 shows the goodness of fit indexes for the different models tested with ESEM. Initial testing of the 3-factor model showed that the overall fit was acceptable: it was good in the Spanish sample and marginally good in the Colombian one. In the latter case the Root mean-square error of approximation (RMSEA) was quite good but the Comparative fit index (CFI) was slightly lower than the standard cut-off values of 0.08 and 0.9, respectively (Bentler, 1990; Browne & Cudeck, 1993).

When configurational invariance was tested, we again found a good RMSEA index (RMSEA= 0.045) and a C.F.I. value close to the 0.9 cut-off value (CFI= 0.88). Nevertheless, as we have stated in the introduction, it is not clear whether these cut-off values are appropriate for ESEM. A more restrictive model (weak factorial invariance) led to a slightly worse fit (RMSEA= 0.049; CFI= 0.85) while the more restrictive model with strong factorial invariance showed a fit that was clearly worse (RMSEA= 0.061; CFI= 0.73), which indicates that the differences in impulsivity between Colombian and Spanish children may be explained by the fact that the instrument lacks strong factorial invariance and not by true differences in impulsivity between the two populations.

	Cognitive	Non-planning
Non-planning	-.24	-
Motor	-.15	.35

Model	χ^2	d.f.	r.m.s.e.a	r.m.s.e.a. 90% CI	cfi
Spain	490.2	250	0.043	0.037 - 0.048	0.910
Colombia	503.37	250	0.045	0.039 - 0.051	0.86
Configurational	1029.87	500	0.045	0.041 - 0.049	0.88
Measurement	1261.8	569	0.049	0.045 - 0.052	0.85
Equal intercepts	1.749	592	0.061	0.058 - 0.065	0.73

Finally, the internal consistency reliabilities of the factor scales are presented in Table 6. As can be observed, the reliabilities are systematically lower in the Colombian sample, although the values are only unacceptable for the *Ic* scale.

Factor	Spanish sample	Colombian sample
Cognitive	.68 (.64; .72)	.59 (.55; .62)
Non-planning	.73 (.69; .76)	.72 (.68; .75)
Motor	.80 (.77; .82)	.74 (.70; .77)

Discussion

The results described above indicate that the factor structure obtained in the BIS 11-c is also found when it is adapted into another culture, at least in terms of configurational invariance. COM has showed that the factor congruence between the Spanish and Colombian versions is high with few items showing inappropriate or borderline congruences. On the other hand ESEM, with the cut-off limitations described above, has showed that both adaptations seem to have the same number of factors and pattern of indicator-factor loadings. Furthermore, the equality of factor loadings (weak factorial invariance) cannot be totally rejected because although authors such as Cheung and Rensvold (2002) consider that a difference in CFI greater than 0.01 is an indicator of worse fit, others such as (Little, 1997) have proposed a value of 0.05, which is higher than the $\Delta CFI= 0.03$ obtained when the weak factorial invariance model is compared with the configurational invariance model in our data. To sum up, the results seem to confirm that both adaptations, which have the same items, have a three-factor structure and show the same relationship between latent variables and scale scores.

Finally the bad fit to the strong invariance model shows that the differences between the Colombian and Spanish sample may not show true differences in impulsivity between the two samples because these differences may also be due to test bias, which under- or overestimates impulsivity levels in one of the samples, procedure differences or differences in sample characteristics. In this sense we have to point one of the limitations of this study that is the incidental nature of the Colombian sample which also may explain the differences between founded between children of both nations. Nevertheless, this is a minor point if, as in this case, we are more interested in obtaining a measure of impulsivity that can be used to predict behaviour problems such as aggression, etc. than in cultural comparisons across nations.

These results seem to indicate that BIS 11-c may be a valuable instrument for assessing impulsivity in children and adolescents, a field in which, as we have stated above, there is a lack of psychometric measurement instruments

Furthermore, the results indicate how important it is to take into account linguistic and cultural issues in test adaptation even when cultures share the same language. In this respect we can see

that the factor structure of the test is congruent after 21 of the 26 items were modified. Of course we do not know what would have happened if we had administered the unmodified Spanish version to the Colombian sample but the results would probably have been worse. Indeed this is a hypothesis that further research will have to prove.

As we have stated in the introduction, this scale showed reliability problems associated to *Icog* when the Spanish version was developed, which was the main reason why new items were added to the first version of the Spanish scale. On the other hand, the factors related to *Ic* have the lowest reliabilities in the different versions of Barrat's Impulsivity Scales (Stanford et al., 2009). Taking this into account we believe that further research is needed to improve the psychometric characteristics of this scale by adding new items at least to the Colombian adaptation. Nevertheless, it should be mentioned that *Im* and *Ihp* showed good or acceptable reliabilities in both versions and that these scales are most related to the dysfunctional aspects of impulsivity, which are the best predictors of aggression, psychopathological problems, addictions, etc.

Now that the factor structure and psychometric properties of two different adaptations of BIS 11-c have been determined, future research will have to show their predictive validity and their convergent validity with other psychometric and behavioural measures of impulsivity.

Finally we should point out that ESEM has some limitations that future research should override. The first is the lack of clear cut-off values to test goodness of fit (see Marsh et al., 2009). The

second is the fact that the ESEM procedure needs to restrict to zero the relationship between the first item and the $p-1$ remaining factors, the relationship between the second item and the $p-2$ remaining factors and so on, chosen them arbitrarily. As a result, depending upon which are the first items in the data file, ESEM might provide a different model fit. In this regard, we believe that is better to choose a set of marker items from previous exploratory studies to define the factors in a semi-restricted factor analysis, as Ferrando & Lorenzo-Seva (2000) proposed, in order to get the best fit possible. In fact, an interesting alternative to the ESEM model may be to extend their model to multigroup analysis. Finally the exploratory nature of ESEM makes it easier to achieve good model fit than with CFA, but we do not believe that this is also the case in multigroup analysis, in which CFA does not test the equivalence between the «minor» loadings for all groups because they are assumed to be zero. ESEM, on the other hand, tests the equivalence of item loadings on all factors. This means that what gives ESEM the advantage over AFC may be its weakness in multigroup analysis. Taking everything into account, although there is no doubt that new methods such as ESEM may be highly valuable in the future as alternatives to classical multigroup CFA, they must be used with caution until all the issues mentioned above have been clarified.

Acknowledgments

This research was supported by a grant from the Spanish Ministry of Education and Science (PSI2008-00236/PSIC).

References

- American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders (DSM-IV-TR). Washington D.C.
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*, 397-438.
- Balluerka, N., Gorostiaga, A., Alonso-Arbiol, I., & Aramburu, M. (2007). La adaptación de instrumentos de medida de una cultura a otra: una perspectiva práctica. *Psicothema, 19*, 124-133.
- Barratt, E.S. (1985). Impulsiveness subtraits: Arousal and information processing. In J.T. Spence & C.E. Izard (Eds.): *Motivation, emotion and personality* (pp. 137-145). New York. North-Holland: Elsevier Science.
- Barratt, E.S., & Slaughter, L. (1998). Defining, measuring and predicting impulsive aggression. A heuristic model. *Behavioral Sciences and the Law, 16*, 285-302.
- Bentler, P.M. (1990). Comparative fit indices in structural models. *Psychological Bulletin, 107*, 238-246.
- Bloxom, B. (1968). Factorial rotation to simple structure and maximum similarity. *Psychometrika, 33*, 237-247.
- Browne, M.W., & Cudek, R. (1993). Alternate ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.): *Testing structural equation models* (pp. 136-162). Newbury Park, CA, Sage.
- Brunas-Walstaff, J., Tilley, A., Verity, M., Ford, S., & Thompson, D. (1997). Functional and dysfunctional impulsivity in children and their relationship to Eysenck's impulsiveness and venturesomeness dimensions. *Personality and Individual Differences, 22*, 19-25.
- Byrne, B.M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema, 20*, 872-882.
- Byrne, B.M., & Campbell, T.L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. *Journal of Cross-cultural Psychology, 30*, 557-576.
- Byrne, B.M., & Stewart, S.M. (2006). The MACS approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural Equation Modeling, 13*, 287-321.
- Byrne, B.M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology, 34*, 155-175.
- Cataldo, M., Nobile, M., Lorusso, M.L., Battaglia, M., & Molteni, M. (2005). Impulsivity in depressed children and adolescents: A comparison between behavioural and neuropsychological data. *Psychiatry Research, 136*, 123-133.
- Chan, W., Ho, R.M., Leung, K., Chan, D.K.S., & Yung, Y-F. (1999). An alternative method for evaluating congruence coefficients with Procrustes rotation: A Bootstrap procedure. *Psychological methods, 4*, 378-402.
- Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255.
- Cosi, S., Canals, J., & Vigil-Colet, A. (2008). Development of an improved version of BIS-11c questionnaire: Relationships with aggression and scholastic performance. Poster presented at the Xth meeting of the Spanish Society for the Study of Individual Differences. Salamanca.
- Cosi, S., Canals, J., Hernández-Martínez, C., & Vigil-Colet, A. (2010). Parent-child agreement in SCARED and its relationship to anxiety symptoms. *Journal of Anxiety Disorders, 24*, 129-133.
- Cosi, S., Morales-Vives, F., Canals, J., Lorenzo-Seva, U., & Vigil-Colet, A. (2008). Functional and dysfunctional impulsivity in childhood and adolescence. *Psychological Reports, 103*, 67-76.
- Cosi, S., Vigil-Colet, A., Canals, J., & Lorenzo-Seva, U. (2008). Psychometric properties of the Spanish Adaptation of the Barratt Impulsiveness Scale-11a for children. *Psychological Reports, 103*, 336-346.
- Dickman, S.J. (1990). Functional and dysfunctional impulsivity, personality and cognitive correlates. *Journal of Personality and Social Personality, 58*, 95-102.
- Eysenck, S.B.G., Easting, G., & Pearson, P.R. (1984). Age norms for impulsiveness, venturesomeness and empathy in children. *Personality and Individual Differences, 5*, 315-321.

- Eysenck, S.B.G., & Eysenck, H.J. (1980). Impulsiveness and venturesomeness in children. *Personality and Individual Differences, 1*, 73-78.
- Ferrando, P.J., & Lorenzo-Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: Some aspects of the problem and some suggestions. *Psicológica, 21*, 301-323.
- Fink, A.D., & McCown, W.G. (1993). Impulsivity in children and adolescents: Measurement, causes and treatment. In W.G. McCown, J.L. Johnson, & M.B. Shure (Eds.). *The impulsive client, theory, research and treatment* (pp. 279-308). Washington: American Psychological Association.
- Fossati, A., Barratt, E.S., Acquarini, E., & Di Ceglie, A. (2002). Psychometric properties of and adolescent version of the Barratt Impulsiveness Scale-11 for a sample of Italian high school students. *Perceptual and Motor Skills, 95*, 621-635.
- Jensen, P.S., Youngstrom, E.A., Steiner, H., Findling, R.I., Meyer, R.E., Malone, R.P., Carlsson, G.A., Coccaro, E.F., Aman, M.G., Blair, J., Dogherty, D., Ferris, C., Flynn, L., Green, E., Hoagwood, K., Hutchinson, J., Laughren, T., Leve, L., Novins, D.K., & Vitello, B. (2007). Consensus report on impulsive aggression as a symptom across diagnostic categories in child psychiatry: Implications for medication studies. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 309-321.
- Gerbing, D.W., Ahadi, S.A., & Patton, J.H. (1987). Toward a conceptualization of impulsivity: Components across the behavioral and self-report domains. *Multivariate Behavioral Research, 22*, 357-379.
- Hambleton, R.K. (2005). Issues, designs and technical guidelines for adapting tests into multiple languages and cultures. In R.K. Hambleton, P.F. Merenda, & C. Spielberger (Eds.). *Adapting educational and psychological tests for cross-cultural Assessment* (pp. 3-38). London: LEA.
- Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31-36.
- Leshem, R., & Glicksohn, J. (2007). The construct of impulsivity revisited. *Personality and Individual Differences, 43*, 681-692.
- Little, T.D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*, 53-76.
- Lorenzo-Seva, U. (2003). A factor simplicity index. *Psychometrika, 68*, 49-60.
- Lorenzo-Seva, U., Kiers, H.A.L., & Ten Berge, J.M.F. (2002). Techniques for oblique factor rotation of two or more loading matrices to a mixture of simple structure and optimal agreement. *British Journal of Mathematical and Statistical Psychology, 55*, 337-360.
- Lorenzo-Seva, U., & ten Berge, J.F.M. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology, 2*, 57-64.
- Lorenzo-Seva, U., & Ferrando, P.J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavioral Research Methods, Instruments and Computers, 38*, 88-91.
- Luengo, M.A., Carrillo-de-la-Pena, M.T., & Otero, J.M. (1991). The components of impulsiveness: A comparison of the I7 impulsiveness questionnaire and the Barratt impulsiveness scale. *Personality and Individual Differences, 12*, 657-667.
- Marsh, H.W., Muthén, B., Asparouhov, A., Lüdtke, O., Robitzsch, A., Morin, A.J.S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling, 16*, 439-476.
- McMurrin, M., Blair, M., & Egan, V. (2002). An investigation of the correlations between aggression, impulsiveness, social problem-solving and alcohol use. *Aggressive Behavior, 28*, 439-445.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543.
- Muris, P., Merckelbach, H., Van Brakel, A., & Mayer, B. (1999). The revised version of the screen for child anxiety related emotional disorders (scared-i): Further evidence for its reliability and validity. *Anxiety, Stress & Coping, 12*, 411-425.
- Patton, J.H., Stanford, M.S., & Barratt, E.S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology, 51*, 768-774.
- Rapee, R.M., Barrett, P.M., Dadds, M.R., & Evans, L. (1994). Reliability of the DSM-III-R childhood anxiety disorders using structured interview: Interrater and parent-child agreement. *Journal of the American Academy of Child and Adolescent Psychiatry, 33*, 984-992.
- Reynolds, B., Ortengren, A., Richards, J.B., & de Wit, H. (2005). Dimensions of impulsive behaviour: Personality and behavioural measures. *Personality and Individual Differences, 40*, 305-315.
- Scarpa, A., & Raine, A. (2000). Violence associated with anger and impulsivity. In J.C. Borod (Ed.). *The neuropsychology of emotion* (pp. 320-339). Oxford: Oxford University Press.
- Silva, F., Martorell, M.C., & Clemente, A. (1987). The I6 questionnaire: Spanish adaptation. *Evaluación Psicológica / Psychological Assessment, 3*, 55-78.
- Stadler, C., & Janke, W. (2003). Concurrent validity of the German version of S.B. Eysenck's impulsiveness questionnaire for children. *Personality and Individual Differences, 35*, 51-58.
- Stanford, M.S., Mathias, C.W., Dougherty, D.M., Lake, S.L., Anderson, N.E., & Patton J.M. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences, 47*, 385-395.
- Summerfeldt, L.J., Hood, K., Antony, M.M., Richter, M.A., & Swinson, R.P. (2004). Impulsivity in obsessive-compulsive disorder: Comparisons with other anxiety disorders and within tic-related subgroups. *Personality and Individual Differences, 36*, 539-553.
- Tucker, L.R. (1951). A method for synthesis of factor analysis studies. *Personnel Research Section Report N° 984*, Washington D.C.: Department of the Army.
- Vigil-Colet, A., & Morales-Vives, F. (2005). How impulsivity is related to intelligence and academic achievement. *The Spanish Journal of Psychology, 8*, 199-204.
- Vigil-Colet, A., Lorenzo-Seva, U., Codomiu-Raga, M.J., y Morales, F. (2005). Factor structure of the aggression questionnaire among different samples and languages. *Aggressive Behavior, 31*, 601-608.
- Vigil-Colet, A., Morales-Vives, F., & Tous, J. (2008). The relationships between functional and dysfunctional impulsivity across samples. *The Spanish Journal of Psychology, 11*, 480-487.
- Willcutt, E.G., & Pennington, B.F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder. Differences by gender and subtype. *Journal of Learning Disabilities, 33*, 179-192.
- Whiteside, S.P., & Lynam, D.R. (2001). The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences, 30*, 669-689.
- Zaparniuk, J., & Taylor, S. (1997). Impulsivity in children and adolescents. In C.D. Webster & M.A. Jackson (Eds.). *Impulsivity: theory, assessment and treatment* (pp. 158-180). New York: Guilford Press.

2 Resultados referentes al test AQ

Los resultados obtenidos con el test AQ se han publicado en: *Chahín-Pinzón, N., Lorenzo-Seva, U., y Vigil-Colet, A. (2012). Características psicométricas de la adaptación colombiana del Cuestionario de Agresividad de Buss y Perry en una muestra de preadolescentes y adolescentes de Bucaramanga. Universitas Psychologica, 11(3), 979-988.*

A continuación se presenta un resumen de los resultados más relevantes, y posteriormente se imprime el artículo tal y como ha sido publicado.

- La adaptación llevó a que se tuvieron que modificar 16 de los 20 ítems para lograr adecuar el instrumento a las características de edad, lingüísticas y culturales propias de la muestra colombiana de niños.
- La estructura original de cuatro factores del AQ compuesta por *agresividad física, agresividad verbal, ira y hostilidad*, presente en versión española para adultos de Vigil-Colet *et al.* (2005), ha sido confirmada en la adaptación para preadolescentes y adolescentes colombianos, con indicadores de ajuste al modelo que informan de una buena adecuación ($RMSEA=.049$, $CFI=.92$, $NFI=.90$).
- La consistencia interna para el total de la muestra fue buena con un $\alpha=.82$. Respecto a las escalas, la que obtuvo el mejor coeficiente fue *agresividad física* con $\alpha=.75$. La escala *hostilidad* ostentó un valor aceptable $\alpha=.62$, pero las otras dos escalas informaron valores por debajo del límite para ser considerados aceptables. El

coeficiente para *agresividad verbal* fue de $\alpha=.58$ (intervalo de confianza entre .52 y .63) y el coeficiente para *ira* de $\alpha=.58$ (intervalo de confianza entre .52 y .63).

- Los resultados de la fiabilidad en los grupos de edad fueron disimiles en comparación a los reportados para la muestra total. En el grupo de 8 y 10 años sólo la escala total con $\alpha=.82$ y la escala *agresividad física* con $\alpha=.80$ obtuvieron buenos coeficientes, los demás valores fueron inaceptables. En el grupo de 11 y 13 años, la escala total con $\alpha=.81$ y la escala *agresividad física* con $\alpha=.73$ presentaron valores adecuados, además la escala *hostilidad* tuvo una consistencia en el límite de lo aceptable con $\alpha=.60$.

- Por otra parte, en el grupo de mayor edad (entre 14 y 16 años), el comportamiento ha sido substancialmente superior. La escala total ha exhibido un coeficiente de $\alpha=.82$, la escala *agresividad física* de $\alpha=.77$, la escala *agresividad verbal* de $\alpha=.66$, la escala *ira* de $\alpha=.63$ y la escala *hostilidad* de $\alpha=.65$. En cambio, en los dos grupos de edad inferior, sólo las puntuaciones de la escala total y *agresividad física* son recomendables con fines clínicos y de investigación.

- El análisis de la varianza realizado para todas las escalas mostró diferencias significativas en la escala *agresividad física* entre los grupos de sexo ($F(1, 488)=29.15; p <.01$), donde los varones puntuaron más alto que las mujeres, con un tamaño del efecto $d=.64$. Por lo demás, la escala *agresividad verbal*, *ira* y *hostilidad* no presentaron diferencias significativas en los grupos de sexo y edad.

Características psicométricas de la adaptación colombiana del Cuestionario de Agresividad de Buss y Perry en una muestra de preadolescentes y adolescentes de Bucaramanga*

Psychometric Properties of Colombian Adaptation of Buss and Perry's Aggression Questionnaire for Teenagers in a Sample of Bucaramanga

Recibido: febrero 17 de 2011 | Revisado: julio 21 de 2011 | Aceptado: septiembre 9 de 2011

NICOLÁS CHAHÍN-PINZÓN**

URBANO LORENZO-SEVA***

ANDREU VIGIL-COLET****

Universidad Rovira i Virgili, Tarragona, España

SICI: 1657-9267(201209)11:3<979;CCABPB>2.0.TX;2-9

Para citar este artículo: Chahín-Pinzón, N., Lorenzo-Seva, U. & Vigil-Colet, A. (2012). Características psicométricas de la adaptación colombiana del Cuestionario de Agresividad de Buss y Perry en una muestra de preadolescentes y adolescentes de Bucaramanga. *Universitas Psychologica*, 11(3), 979-988.

* Agradecimientos: Esta investigación ha estado financiada por una beca del Ministerio Español de Educación y Ciencia (PSI2008-00236/PSIC).

** Doctorando programa de doctorado en Evaluación y Medida de la Conducta de la Universidad Rovira i Virgili. E-mail: nicocha@yahoo.com

*** Catedrático del Área de Metodología de las Ciencias del Comportamiento. Departamento de Psicología. E-mail: urbano.lorenzo@urv.cat. ResearcherID: Lorenzo-Seva, U. G-4228-2011.

**** Catedrático del Área de Metodología de las Ciencias del Comportamiento. Departamento de Psicología. E-mail: andreu.vigil@urv.cat. ResearcherID: Vigil-Colet, A. B-8650-2011.

RESUMEN

Este estudio presenta la adaptación colombiana del Cuestionario de Agresividad de Buss y Perry para preadolescentes y adolescentes, desarrollado a partir de la versión española abreviada de 20 ítems. Dicha adaptación se administró a una muestra de 535 niños (269 niños y 266 niñas) con un rango de edades de 8-16 años, pertenecientes a tres colegios del área metropolitana de Bucaramanga. Los resultados del análisis factorial confirmatorio indican un buen ajuste al modelo de cuatro factores que ha sido descrita para otros idiomas y adaptaciones. El test presentó una fiabilidad satisfactoria para la escala total ($\alpha = 0.82$) y para la escala de agresividad física ($\alpha = 0.75$), mientras que para las demás escalas varía en función de la edad, no siendo recomendable su aplicación en edades inferiores a los 12 años. Al analizar los efectos del sexo sobre la agresividad, se verificó que dicho efecto se debía exclusivamente a las diferencias en agresividad física. Finalmente, se pone de manifiesto la importancia de considerar los aspectos culturales y lingüísticos en la adaptación de un instrumento, incluso, como en este caso, cuando se trata del mismo idioma.

Palabras clave autores:

Agresividad, adaptación de escalas, adolescencia.

Palabras clave descriptores:

Adaptación de test, pruebas psicológicas, investigación cuantitativa.

ABSTRACT

This study presents the Colombian adaptation of the Buss and Perry's Aggression Questionnaire for preadolescents and adolescents. The test was administered to a sample of 535 children (269 boys and 266 girls) with an age range of 8-16 years old, belonging to three schools of Bucaramanga. Confirmatory factor analysis showed a good fit to the four-factor model and reliability was satisfactory. The reliability was good for overall scale ($\alpha = 0.82$) and physical aggression scale ($\alpha = 0.75$) while it was sufficient or poor for the other scales depending upon the age group, not being suitable for ages under 12 years-old. Referring to sex effects, we found them only for physical aggression. Finally, this study places particular emphasis on the importance of linguistic and cultural aspects in test adaptation, even when both cultures share the same language.

Key words authors:

Aggression, Scale Adaptation, Adolescence.

Key words plus:

Test Adaptation, Psychological Tests, Quantitative Research.

Introducción

El Cuestionario de Agresividad (AQ) de Buss y Perry (1992) es, en la actualidad, uno de los instrumentos psicométricos más utilizados para evaluar e investigar la conducta agresiva. El AQ se desarrolló a partir del Hostility Inventory de Buss y Durkee (1957), que comprendía siete escalas y 75 ítems. Pero sus limitaciones psicométricas llevaron a Buss y Perry a una versión más refinada, compuesta 29 ítems, con una estructura factorial que comprende cuatro escalas de agresividad: física y verbal, que representan los componentes instrumental y motor; ira, el componente afectivo; y hostilidad, el componente cognitivo de la agresión. Su estructura factorial ha sido estudiada también en diversos idiomas y culturas, como es el caso, entre otras, de la versión japonesa (Nakano, 2001; Ramírez, Andreu & Fujihara, 2001), eslovaca (Lovas & Trenkova, 1996), alemana (Von Collani & Werner, 2005), holandesa (Meesters, Muris, Bosma, Schouten & Beuving, 1996), francesa (Pfister & Masse, 2001), italiana (Fossati, Maffei, Acquarini & Di Ceglie, 2003) y española (Andreu, Peña & Graña, 2002; García-León et al., 2002).

A pesar de la elevada coincidencia de numerosos estudios en la solución de los cuatro factores, diversos investigadores han encontrado problemas en el ajuste de algunos ítems, especialmente los revertidos y los pertenecientes a la Escala de Hostilidad (Bernstein & Gesn, 1997; Harris, 1995; Meesters et al., 1996). Teniendo en cuenta que estos ajustes afectaban a distintos ítems en las diversas adaptaciones del cuestionario, se desarrolló una versión del AQ para el contexto español que eliminaba aquellos ítems que presentaban una baja congruencia a través de las distintas culturas (Vigil-Colet, Lorenzo-Seva, Codorniu-Raga & Morales, 2005). Como resultado, se obtuvo una nueva versión reducida para adultos de 20 ítems que presenta un ajuste satisfactorio cuando ha sido evaluada mediante análisis factorial confirmatorio. Además, por el procedimiento que se ha llevado a cabo se garantiza una mayor validez transcultural que permite llevar a cabo comparaciones de los resultados entre diferentes culturas (Morales-Vives, Codorniu-Raga & Vigil-Colet, 2005).

Un aspecto muy importante para tener presente acerca del AQ, es que a pesar de que fue desarrollado inicialmente para ser administrado en adultos, ha señalado su adecuación en poblaciones jóvenes. Es el caso de la adaptación para adolescentes y preadolescentes desarrollada por Santisteban, Alvarado y Recio (2007) a partir de la versión original del AQ, que ha mostrado una adecuada consistencia interna y un buen ajuste al modelo de cuatro factores. Lo que indica que este puede ser útil para obtener estimaciones fiables y válidas de la agresividad en estos rangos de edad (Santisteban & Alvarado, 2009). Por otra parte, también se ha encontrado en Asia que la versión abreviada del AQ de 12 ítems, presenta características similares en adolescentes (Ang, 2007).

En América Latina, el AQ ha tenido solo dos adaptaciones, ambas para adultos. En la primera, desarrollada en Colombia por Castrillón, Ortiz y Vieco (2004), se encontraron cinco factores, añadiendo el factor de No Agresión a los cuatro descritos originalmente; además en el factor de hostilidad se encontraron diferencias respecto a estudios anteriores. Los propios autores puntualizaron que las diferencias encontradas podían deberse a efectos culturales. Por otra parte, hay que señalar que este estudio no partió la versión original del AQ, ya que a esta se le agregaron 11 ítems. En la segunda adaptación, realizada en El Salvador, se encontró la misma estructura factorial del AQ original, aunque los factores presentaron ligeras diferencias en su composición y dos de los ítems debieron ser eliminados por su baja comunalidad (Sierra & Gutiérrez, 2007).

Debido a que en Latinoamérica no existe en la actualidad una versión del AQ para preadolescentes y adolescentes, se decidió realizar en esta investigación una adaptación para dicha población, a partir de la versión reducida para adultos desarrollada en España (Morales-Vives et al., 2005). Se consideró que esta tiene la ventaja adicional de haber sido desarrollada teniendo en cuenta características transculturales, que se encuentra libre de sesgos debidos al sexo y que su aplicación es mucho más rápida ya que posee un número menor de ítems (Condon, Morales-

CARACTERÍSTICAS PSICOMÉTRICAS DE LA ADAPTACIÓN COLOMBIANA DEL CUESTIONARIO DE AGRESIVIDAD DE BUSS Y PERRY

Vives, Ferrando & Vigil-Colet, 2006; Vigil-Colet et al., 2005).

Sin embargo, realizar una adaptación del AQ desde dos contextos tan distintos, como lo son España y Colombia, requiere no perder de vista algunos aspectos previos. Aunque la adaptación de pruebas por lo general implica traducción de un idioma a otro, esta va mucho más allá de una simple traducción, ya que implica el análisis de las diferencias culturales y de las variedades lingüísticas de la cultura original y de aquella a la que se pretende adaptar el instrumento (Hambleton, 2005). Este análisis de las diferencias culturales y lingüísticas es un aspecto clave en varias de las directrices propuestas por la Comisión Internacional de Tests (ITC) para la adaptación de tests en distintas culturas (Hambleton, 1994; van de Vijver & Hambleton, 1996). No obstante, es muy frecuente el caso en el que un instrumento adaptado o desarrollado en España, por el hecho de estar en idioma español se asuma a priori que sus propiedades psicométricas van a ser las mismas al ser aplicado en cualquier otro país de habla hispana. Esta idea no considera que las variantes lingüísticas y las diferencias culturales afectan el funcionamiento del instrumento de medida. Es así como una misma expresión puede implicar un nivel distinto del rasgo que se pretende medir en función de cada país. Por ejemplo, “amenazar a alguien” no implica en España una alta agresividad, mientras que en Colombia significa un nivel muy elevado de agresividad, en el que está en riesgo la vida del individuo amenazado. Por lo argumentado, se considera que la aplicación del AQ en población preadolescente y adolescente, requiere previamente de un proceso serio de adaptación específico, en lugar de aplicar directamente la versión española.

Por tanto, el objetivo de este estudio es el desarrollo de una versión reducida del AQ ajustada a las características de la población preadolescente y adolescente colombiana, a partir de la realizada en el contexto español (Morales-Vives et al., 2005). En esta nueva adaptación, se estudiarán sus propiedades psicométricas y su estructura factorial. Por otra parte, y considerando la especial importancia de la evaluación de la agresividad en estas edades,

se pretende igualmente analizar sus propiedades en los diferentes grupos de edad.

Método

Participantes

A partir de una muestra inicial de 616 sujetos se eliminaron aquellos casos en los que faltaba la respuesta a algún ítem, obteniéndose una muestra final constituida por 535 sujetos. Las edades de la muestra oscilan entre 8 y 16 años ($M = 12.74$, $DE = 2.111$), donde 266 (49.72 %) son mujeres ($M = 12.94$, $DE = 2.095$) y 269 (50.28 %) hombres ($M = 12.54$, $DE = 2.112$). La diferencia entre la edad promedio de hombres y mujeres resultó significativa ($t = 2.224$, $p = 0.027$), aunque el tamaño del efecto extremadamente bajo ($d = 0.09$) permite considerar a ambos grupos prácticamente equivalentes en edad. La muestra está compuesta por estudiantes de los grados 4° al 11°. Los participantes pertenecen a tres colegios situados en el área metropolitana de Bucaramanga (Colombia), Centro Piloto Simón Bolívar ($N = 240$), Instituto INEM Custodio García Rovira ($N = 114$) y Colegio Gimnasio Superior ($N = 181$), los dos primeros de carácter oficial y el último, privado. La distribución de la muestra por rango de edades, sexo y grado puede apreciarse en la Tabla 1.

Procedimiento

Inicialmente se procedió a contactar con los rectores, coordinadores de jornada, directores de curso y psico-orientadores de los tres centros educativos seleccionados, con el fin de concertar las entrevistas para explicar los objetivos y procedimientos que se llevarían a cabo, y obtener su autorización y la de los padres. Una vez obtenidos los permisos respectivos, en presencia de cada profesor se informó a los estudiantes sobre la actividad que se desarrollaría. La participación fue totalmente voluntaria. El cuestionario fue aplicado por un psicólogo acompañado por un profesor, dentro del horario de clase, en grupos de 30 personas, en los salones donde habitualmente se tienen las actividades escolares. Las instrucciones, condiciones de

TABLA 1

Distribución de la muestra por edades, sexo y grado

	Edad	Grado*								Total	
		Col	4	5	6	7	8	9	10		11
		Es	Primaria	Secundaria-Obligatoria			Bachillerato				
Mujeres	8-10		13	15	4	0	0	0	0	0	32
	11-13		4	7	57	11	21	21	0	0	121
	14-16		0	0	1	1	6	40	28	37	113
	Total		17	22	62	12	27	61	28	37	266
Varones	8-10		20	27	7	0	0	0	0	0	54
	11-13		6	10	56	18	13	13	0	0	116
	14-16		0	0	0	1	17	32	31	18	99
	Total		26	37	63	19	30	45	31	18	269
Total	8-10		33	42	11	0	0	0	0	0	86
	11-13		10	17	113	29	34	34	0	0	237
	14-16		0	0	1	2	23	72	59	55	212
	Total		43	59	125	31	57	106	59	55	535

*Se dan los grados para Colombia (Col) y su equivalencia para el sistema educativo de España (Es).

Fuente: elaboración propia.

aplicación y manejo de la información, garantizaban la confidencialidad y el anonimato de todos los datos. El tiempo para resolver el cuestionario osciló entre los 12 y 15 minutos.

Instrumentos

Versión reducida del cuestionario AQ de Buss y Perry en español (Vigil-Colet et al., 2005), compuesto por 20 ítems en escala tipo Likert con cinco alternativas (1 = *Muy rara vez*; 5 = *Muy frecuentemente*). Posee cuatro escalas: agresividad física (7 ítems), agresividad verbal (4 ítems), ira (4 ítems) y hostilidad (5 ítems). La fiabilidad de las mismas es de $\alpha = 0.88$; 0.71; 0.68 y 0.65, respectivamente, mientras que para la escala total es de $\alpha = 0.87$ (Morales-Vives et al., 2005).

Con el fin de adaptar dicho instrumento para su administración en Colombia, un psicólogo colombiano y dos españoles con experiencia en la adaptación de cuestionarios analizaron su adecuación cultural y lingüística. Dicho análisis implicó modificar en mayor o menor medida 16 de los 20 ítems, bien por diferencias culturales o lingüísticas, o para adaptar su contenido a este tipo de población. Así por ejemplo, se modificaron expresiones como: "me

incitan" por "me provocan", "recurrir a la violencia" por "pelear" o "me enfado rápidamente" por "me da rabia fácilmente", entre otras.

El cuestionario resultante en una segunda fase fue evaluado por ocho licenciados colombianos dedicados a la enseñanza, en relación a la facilidad de comprensión de los ítems por parte de sus alumnos. Los conceptos de los docentes llevaron a unos cuantos cambios ligeros en algunos de los ítems, por ejemplo se sustituyó "golpear" por "pegar" y "provocar" por "molestar". Finalizada esta fase, se obtuvo la versión definitiva del AQ para ser administrada en la muestra seleccionada de preadolescentes y adolescentes. La Tabla 2 muestra esta versión junto a la original española.

Análisis de datos

Partiendo de que el AQ posee una estructura factorial que ha sido ampliamente replicada en numerosos idiomas y culturas, se ha considerado que la aproximación más adecuada para analizar la estructura factorial de esta nueva versión sea de tipo confirmatorio. Para tal fin, se ha adoptando la estructura de cuatro factores de las versiones española y anglosajona como hipótesis inicial.

CARACTERÍSTICAS PSICOMÉTRICAS DE LA ADAPTACIÓN COLOMBIANA DEL CUESTIONARIO DE AGRESIVIDAD DE BUSS Y PERRY

TABLA 2

Ítems del AQ de la versión española y de la adaptación colombiana para preadolescentes y adolescentes

Escala	Versión española	Versión colombiana
Agresividad Física	De vez en cuando no puedo controlar el impulso de golpear a otra persona.	De vez en cuando no puedo controlar el impulso de golpear a otra persona.
	Si se me provoca lo suficiente, puedo golpear a otra persona.	Si me molestan mucho, puedo llegar a pegarle a otra persona.
	Si alguien me golpea, le respondo golpeándole también.	Si me pegan, yo devuelvo el golpe.
	Si tengo que recurrir a la violencia para proteger mis derechos, lo hago.	Si tengo que pelear para defender mis derechos, lo hago.
	Hay gente que me incita hasta tal punto que llegamos a pegarnos.	Hay personas que me molestan tanto que terminamos pegándonos.
	He amenazado a gente que conozco.	He intimidado a personas que conozco.
Agresividad Verbal	He llegado a estar tan furioso que rompía cosas.	Cuando me han "sacado la piedra" he dañado cosas.
	Cuando no estoy de acuerdo con mis amigos, discuto abiertamente con ellos.	Cuando no estoy de acuerdo con mis amigos, discuto con ellos.
	A menudo no estoy de acuerdo con la gente.	Con frecuencia no estoy de acuerdo con la gente.
Ira	Cuando la gente no está de acuerdo conmigo, no puedo evitar discutir con ellos.	Cuando los demás no están de acuerdo conmigo, no puedo evitar discutir con ellos.
	Mis amigos dicen que discuto mucho.	Mis amigos dicen que discuto mucho.
	Me enfado rápidamente, pero se me pasa enseguida.	Me da rabia fácilmente, pero se me pasa rápido.
Hostilidad	Cuando estoy frustrado, muestro el enfado que tengo.	Cuando tengo rabia, no la disimulo.
	Algunas veces me siento tan enfadado como si estuviera a punto de estallar.	Algunas veces tengo tanta rabia que me siento como si estuviera a punto de explotar.
	Algunas veces pierdo los estribos sin razón.	Algunas veces se me "salta la piedra" sin razón.
	A veces soy bastante envidioso.	A veces soy bastante envidioso.
	Me pregunto por qué algunas veces me siento tan resentido por algunas cosas.	Algunas veces me pregunto por qué me siento tan resentido por algunas cosas.
Hostilidad	Sé que mis amigos me critican a mis espaldas.	Sé que mis amigos me critican a mis espaldas.
	Cuando la gente se muestra especialmente amigable, me pregunto qué querrán.	Cuando las personas se muestran muy amigables, me pregunto qué es lo que quieren.
	Algunas ocasiones siento que la gente se está riendo de mí a mis espaldas.	A veces siento que la gente se ríe de mí a mis espaldas.

Fuente: elaboración propia.

Los índices de ajuste utilizados para evaluar el análisis factorial confirmatorio (AFC) serán los propuestos por Hu y Bentler (1999), como lo son el índice de ajuste comparativo (CFI), el índice de ajuste no normativo (NFI) y la raíz cuadrada media de error de aproximación (RMSEA). Aunque no existe un consenso absoluto, se consideran habitualmente aceptables valores de corte iguales o superiores a 0.9 para los dos primeros, mientras que para el RMSEA son aceptables valores inferiores a 0.08 y excelentes los iguales o inferiores a 0.05 (Bentler, 1990; Fan & Sivo, 2007). El AFC se llevó

a cabo utilizando el software Mplus 5.1. (Muthen & Muthen, 2008) sobre la matriz de correlaciones policóricas, debido a que las respuestas a los ítems pueden considerarse como pertenecientes a una escala de medida ordinal.

En caso que la estructura factorial se confirmara, se analizaría la fiabilidad de las escalas, así como la puntuación total mediante el índice alfa de Cronbach. Este análisis se llevaría a cabo para toda la muestra y para los distintos grupos de edad, con el fin de determinar a partir de qué edad puede considerarse apropiada la utilización del AQ

en esta versión. Estos análisis junto con el cálculo de los estadísticos descriptivos y el análisis de los efectos de edad y sexo, se llevaron a cabo utilizando el software SPSS 17.0.

Resultados

El AFC planteado para la estructura de cuatro factores proporcionó los siguientes índices de ajuste del modelo: CFI = 0.92, NFI = 0.90 y RMSEA = 0.049 (I.C. al 90 %: 0.044 - 0.057), con una probabilidad $p = 0.349$ de que el valor poblacional de RMSEA sea inferior a 0.05. De este modo, y teniendo en cuenta los puntos de corte de estos indicadores expuestos con anterioridad, puede considerarse que todos ellos muestran un ajuste más que aceptable del modelo, siendo excelente en el caso del RMSEA. Por consiguiente, este análisis indica la adecuación de la adaptación colombiana del AQ para preadolescentes y adolescentes a la estructura de cuatro factores, replicándose los factores de Agresividad Física, Agresividad Verbal, Ira y Hostilidad descritos tanto para la versión original como para las subsiguientes adaptaciones a otros idiomas y versiones.

Una vez establecida la estructura factorial, se calcularon los estadísticos descriptivos de las es-

calas. La Tabla 3 muestra los estadísticos de las cuatro escalas para los grupos de sexo y para los tres grupos de edad.

El análisis multivariado de la varianza utilizando los grupos de edad y el sexo como factores, mostró la inexistencia de efectos tanto de la variable edad como de la interacción Edad x Sexo sobre las escalas del AQ. En cambio, el análisis sí mostró efectos significativos de la variable sexo sobre las escalas del AQ ($F_{(4,485)} = 9.72$; $p < 0.01$). Y al analizar los efectos del sexo sobre cada una de las cuatro escalas a nivel univariado, se verificó que dicho efecto se debía exclusivamente a las diferencias de sexo en agresividad física ($F_{(1,488)} = 29.15$; $p < 0.01$), presentando los jóvenes puntuaciones superiores ($d = 0.64$).

La Tabla 4 muestra las fiabilidades obtenidas para el total de la muestra y para los distintos grupos de edad. Tal y como se puede comprobar, solo la escala de Agresividad Física y la escala total presentan coeficientes de fiabilidad adecuados en todos los grupos de edad. Por su parte, en el grupo de edad mayor (14 - 16 años) las escalas de Agresividad Verbal, Ira y Hostilidad presentan coeficientes que podrían considerarse como marginalmente aceptables, puesto que el valor $\alpha = 0.70$, generalmente considerado como aceptable, se encuentra

TABLA 3
Estadísticos descriptivos del AQ de la adaptación colombiana desglosados por sexo y edad

		Varones		Mujeres					
		Media	DE	Media	DE				
Escala	Física	17.98	5.6	14.57	5.1				
	Verbal	9.19	3	9.01	3.3				
	Ira	10.5	3.6	10.7	3.8				
	Hostilidad	12.80	4.1	12.85	4.1				
	Total	50.54	11.7	47.14	12.42				
		Grupos de edad							
		8 a 10 años		11 a 13 años		14 a 16 años		Total	
		Media	DE	Media	DE	Media	DE	Media	DE
Escala	Física	15.17	6.2	16.1	5.4	16.69	5.6	16.26	5.6
	Verbal	8.69	3.1	8.94	2.9	9.38	3.4	9.12	3.2
	Ira	10.14	3.6	10.43	3.6	10.92	3.7	10.62	3.6
	Hostilidad	13.14	4	13.03	4.1	12.54	4.2	12.81	4.1
	Total	47.16	12.8	48.52	12.1	49.54	12.1	48.83	12.1

Fuente: elaboración propia.

CARACTERÍSTICAS PSICOMÉTRICAS DE LA ADAPTACIÓN COLOMBIANA DEL CUESTIONARIO DE AGRESIVIDAD DE BUSS Y PERRY

TABLA 4

Coefficiente α Cronbach e intervalos de confianza al 95% para las escalas de la adaptación colombiana del AQ para los grupos de edad

		Grupos de edad			
		8 a 10 años	11 a 13 años	14 a 16 años	Total
Escala	Física	0.80 (0.72-0.86)	0.73 (0.66-0.78)	0.77 (0.72-0.81)	0.75 (0.72-0.78)
	Verbal	0.48 (0.27-0.64)	0.51 (0.39-0.60)	0.66 (0.59-0.73)	0.58 (0.52-0.63)
	Ira	0.36 (0.11-0.52)	0.58 (0.46-0.64)	0.63 (0.55-0.70)	0.58 (0.52-0.63)
	Hostilidad	0.61 (0.47-0.73)	0.60 (0.51-0.68)	0.65 (0.57-0.72)	0.62 (0.57-0.65)
	Total	0.82 (0.75-0.88)	0.81 (0.77-0.85)	0.83 (0.79-0.86)	0.82 (0.80-0.84)

Fuente: elaboración propia.

dentro cada uno de los intervalos de confianza de estas escalas.

Discusión

Los resultados presentados en este estudio señalan la adecuación del AQ en su adaptación colombiana para preadolescentes y adolescentes, a la estructura factorial descrita para otros idiomas y adaptaciones tanto para preadolescentes y adolescentes (Ang, 2007; Santisteban & Alvarado, 2009), como para adultos (Bryant & Smith, 2001; Buss & Perry, 1992; Vigil-Colet et al., 2005). Además, el buen ajuste del modelo logrado después de las numerosas modificaciones llevadas a cabo sobre los ítems de la versión española, va en línea con otros trabajos que han demostrado la importancia de no perder de vista las peculiaridades lingüísticas y culturales del entorno en que se aplican los cuestionarios (Chahín, Cosi, Lorenzo-Seva & Vigil-Colet, 2010; Hambleton, 2005). De este modo, esta versión puede convertirse en un instrumento valioso, útil y práctico a la hora de evaluar los niveles de agresividad con fines diagnósticos o preventivos.

Con referencia a la fiabilidad, la presente adaptación ha mostrado que la escala total y la agresividad física poseen coeficientes adecuados para toda la muestra y para cada uno de los grupos de edad. Sin embargo, hay que considerar que en el grupo de menor edad (8 - 10 años), solo la escala de total y la de agresividad física son aptas para ser aplicadas, ya que las otras exhiben unos valores relativamente bajos. Por otra parte, en el grupo de mayor edad (14 -16 años) la confiabilidad de las es-

calas de agresividad verbal, ira y hostilidad tiende a mejorar, presentando coeficientes marginalmente aceptables. Cabe señalar que estos valores en los distintos grupos de edad no son exclusivos de la adaptación colombiana, pues resultados similares con muestras de adolescentes y preadolescentes han sido reportados, tanto en versiones largas adaptadas a esta población (Santisteban & Alvarado, 2009), como en versiones reducidas (Ang, 2007). Es preciso resaltar que la menor confiabilidad de las escalas de agresividad verbal, ira e y hostilidad ha sido una constante desde que Buss y Perry (1992) desarrollaron la primera versión del AQ, en la que reportaron confiabilidades alrededor de $\alpha = 0.7$ para dichas escalas, mientras que para la escala total y la de agresividad física se situó en $\alpha = 0.8$, obteniéndose los mismos valores para la versión reducida española a partir de la cual se ha desarrollado la adaptación colombiana (Morales-Vives et al., 2005). Algunos estudios incluso han señalado confiabilidades en adultos notablemente inferiores a las reportadas por nosotros en adolescentes y preadolescentes. Así Meesters et al. (1996) y García-León et al. (2002) reportaron una $\alpha = 0.5$ para la escala de agresividad verbal en la adaptación holandesa y española, respectivamente. Todo ello indica que desde el punto de vista de la confiabilidad, la adaptación colombiana puede considerarse equivalente a las existentes en otros idiomas cuando se aplica a población adolescente y, ligeramente inferior en algunas escalas en población preadolescente.

No obstante lo anterior, diversos estudios han señalado que en general los datos obtenidos mediante autoinformes en niños a partir de 8 años, a

pesar de presentar fiabilidades relativamente bajas, suelen ser mejores indicadores de su comportamiento que aquellos respondidos por sus padres y profesores (Achenbach, McConaughy & Howell, 1987; Cosi, Canals, Hernández-Martínez & Vigil-Colet, 2010; Muris, Meesters & Schouten, 2002). Es por esto que a pesar de la baja fiabilidad de las escalas de agresividad verbal, ira y hostilidad, no se descartaría la utilización de todas las escalas del cuestionario entre los 8 y 14 años. En último término, para estas edades podría complementarse la información proporcionada por el AQ con otros cuestionarios en español, que son los aplicados a sus profesores o padres, como el Test de Agresividad Proactiva/Reactiva (Cosi, Vigil-Colet & Canals, 2009).

Al analizar las diferencias de agresividad entre sexos, se ha verificado también que esta adaptación colombiana del AQ para preadolescentes y adolescentes es sensible a las diferencias en agresividad física. Esto mismo ha sido reportado en numerosas investigaciones, en el sentido que los varones presentan niveles de agresividad física más altos que las mujeres, en un amplio rango de edades que va desde la adolescencia hasta la tercera edad (Archer, 2004; Morales-Vives & Vigil-Colet, 2010). Este resultado aporta una primera evidencia de validez de constructo para la presente versión. Cabe resaltar que el tamaño de efecto encontrado para esta diferencia ($d = 0.64$) es similar al reportado en otras investigaciones en población adolescente como las de Santisteban y Alvarado (2009) o Vigil-Colet et al. (2008) que presentaron tamaños del efecto de $d = 0.69$ y $d = 0.47$, respectivamente. De este modo, las diferencias en agresividad física en esta muestra colombiana pueden considerarse equivalentes a las encontradas en adolescentes españoles.

A partir de los resultados aquí presentados, se abre una primera línea de acción para futuras investigaciones que debe aportar evidencia sobre el grado de validez convergente del AQ en estas edades con otros cuestionarios y dimensiones de personalidad que han sido frecuentemente relacionadas con la agresividad, así como sobre su validez predictiva en relación con los comportamientos violentos. Del mismo modo, una segunda línea debe permitir en el futuro llevar a cabo comparaciones

interculturales sobre los niveles de agresividad, donde se pueda establecer si las distintas adaptaciones están midiendo el constructo de la misma manera. Este es un aspecto clave a la hora de llevar a cabo comparaciones transculturales y se relaciona con el establecimiento de una serie de igualdades jerárquicas entre las distintas adaptaciones (Byrne, 2008; Byrne & Watkins, 2003). Estas igualdades se refieren al grado en que las distintas adaptaciones presentan los mismos factores e ítems (invarianza configuracional), al grado en que la relación entre las variables latentes y los ítems es equivalente (igualdad de saturaciones factoriales o invarianza factorial débil) y, finalmente, la invarianza factorial fuerte en el caso de igualdad de los interceptos en la relación entre las variables latentes y los ítems (Meredith, 1993). Tan solo el establecimiento de estas tres igualdades garantiza que las posibles diferencias entre, por ejemplo, la versión española y la colombiana, sean debidas a auténticas diferencias en agresividad y no a un funcionamiento diferencial de los ítems. Con referencia a lo expuesto anteriormente, el presente trabajo ha establecido que la estructura factorial y la composición de los factores del AQ es la misma en las versiones colombiana y española, aportando evidencia sobre la invarianza configuracional. Con todo lo señalado, si se desea llevar a cabo en el futuro comparaciones entre las puntuaciones del AQ entre ambos países, sería necesario utilizar muestras de edad y nivel socio-cultural equivalentes.

Referencias

- Achenbach, T. M., McConaughy, S. H. & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232.
- Andreu, J. M., Peña, M. E. & Graña, J. L. (2002). Adaptación psicométrica de la versión española del Cuestionario de Agresión. *Psicothema*, 14, 476-482.
- Ang, R. (2007). Factor structure of the 12-item Aggression Questionnaire: Further evidence from Asian adolescent samples. *Journal of Adolescence*, 30, 671-685.

CARACTERÍSTICAS PSICOMÉTRICAS DE LA ADAPTACIÓN COLOMBIANA DEL CUESTIONARIO DE AGRESIVIDAD DE BUSS Y PERRY

- Archer, J. (2004). Sex differences in real-word settings: A meta-analytic review. *Review of General Psychology*, 8, 291-332.
- Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, 107, 238-246.
- Bernstein, I. H. & Gesn, P. R. (1997). On the dimensionality of the Buss/Perry Aggression Questionnaire. *Behavioral Research and Therapy*, 35, 563-568.
- Bryant, F. B. & Smith, B. D. (2001). Refining the architecture of aggression: A measurement model for the Buss-Perry Aggression Questionnaire. *Journal of Research on Personality*, 35, 138-167.
- Buss, A. H. & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology*, 21, 343-349.
- Buss, A. H. & Perry, M. P. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63, 452-459.
- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, 20, 872-882.
- Byrne, B. M. & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, 34, 155-175.
- Castrillón, D. A., Ortiz, P. A. & Vieco, F. (2004). Cualidades paramétricas del Cuestionario de Agresión (AQ) de Buss y Perry en estudiantes universitarios de la ciudad de Medellín (Colombia). *Revista Facultad Nacional de Salud Pública*, 22, 49-61.
- Chahín, N., Cosi, S., Lorenzo-Seva, U. & Vigil-Colet, A. (2010). Stability of the factorial structure of Barrat's Impulsivity Scales for children across cultures: A comparison between Spain and Colombia. *Psicothema*, 22, 983-989.
- Condon, L., Morales-Vives, F., Ferrando, P. J. & Vigil-Colet, A. (2006). Sex differences in the full and reduced versions of the Aggression Questionnaire: A question of differential item functioning? *European Journal of Psychological Assessment*, 22, 92-97.
- Cosi, A., Vigil-Colet, A. & Canals, J. (2009). Desarrollo del Cuestionario de Agresividad proactiva/reactiva para profesores: estructura factorial y propiedades psicométricas. *Psicothema*, 21, 159-164.
- Cosi, A., Canals, J., Hernández-Martínez, C. & Vigil-Colet, A. (2010). Parent-child agreement in SCARED and its relationship to anxiety symptoms. *Journal of Anxiety Disorders*, 24, 129-133.
- Fan, X. & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, 42, 509-529.
- Fossati, A., Maffei, C., Acquarini, E. & Di Ceglie, A. (2003). Multigroup confirmatory component and factor analyses of the Italian version of the Aggression Questionnaire. *European Journal of Psychological Assessment*, 19, 54-65.
- García-León, A., Reyes, G. A., Vila, J., Pérez, N., Robles, H. & Ramos, M. M. (2002). The Aggression Questionnaire: A validation study in student samples. *The Spanish Journal of Psychology*, 5, 45-53.
- Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10, 229-244.
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda & C. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). London: L.E.A.
- Harris, J. A. (1995). Confirmatory factor analysis of the Aggression Questionnaire. *Behavioral Research and Therapy*, 8, 991-993.
- Hu, L. & Bentler, P. M. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Lovas, L. & Trenkova, S. (1996). Aggression and perception of an incident. *Studia Psychologica*, 38, 265-270.
- Meesters, C., Muris, P., Bosma, H., Schouten, E. & Beuving, S. (1996). Psychometric evaluation of the Dutch version of the Aggression Questionnaire. *Behavioral Research and Therapy*, 34, 839-843.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525-543.
- Morales-Vives, F., Codorniu-Raga, M. J. & Vigil-Colet, A. (2005). Psychometric properties of the reduced versions of Buss and Perry's Aggression Questionnaire. *Psicothema*, 17, 96-100.

- Morales-Vives, F. & Vigil-Colet, A. (2010). Are there sex differences in physical aggression in the elderly? *Personality and Individual Differences, 49*, 659-662.
- Muris, P., Meesters, C. & Schouten, E. (2002). A brief questionnaire of DSM-IV-Defined anxiety and depression symptoms among children. *Clinical Psychology and Psychotherapy, 9*, 430-442.
- Muthen, L. K. & Muthen, B. O. (2008). Mplus (Version 5.1). Los Angeles, CA: Muthen & Muthen.
- Nakano, K. (2001). Psychometric evaluation on the Japanese adaptation of the Aggression Questionnaire. *Behavioral Research and Therapy, 39*, 853-858.
- Pfister, R. & Masse, P. (2001, noviembre). *Le Questionnaire d'Aggressivité: validation et équivalence transculturelle*. IXème Congrès International des Chercheurs en Activités Physiques et Sportives, Valence, France.
- Ramírez, J., Andreu, J. M. & Fujihara, T. (2001). Cultural and sex differences in aggression: A comparison between Japanese and Spanish students using two different inventories. *Aggressive Behavior, 27*, 313-322.
- Santisteban, C. & Alvarado, J. M. (2009). The Aggression Questionnaire for Spanish preadolescents and adolescents: PA-AR. *The Spanish Journal of Psychology, 12*, 320-326.
- Santisteban, C., Alvarado, J. M. & Recio, P. (2007). Evaluation of a Spanish version of the Buss and Perry Aggression Questionnaire: Some personal and situational factors related to the aggression scores of young subjects. *Personality and Individual Differences, 42*, 1453-1462.
- Sierra, J. C. & Gutiérrez, J. R. (2007). Validación de la versión española del Cuestionario de Agresión de Buss-Perry en estudiantes universitarios salvadoreños. *Psicología y Salud, 107*, 103-113.
- van de Vijver, F. J. R. & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist, 1*, 89-99.
- Vigil-Colet, A., Lorenzo-Seva, U., Codorniu-Raga, M. J. & Morales, F. (2005). Factor structure of the Aggression Questionnaire among different samples and languages. *Aggressive Behavior, 31*, 601-608.
- Von Collani, G. & Werner, R. (2005). Self-related and motivational constructs as determinants of aggression. An analysis and validation of a German version of the Buss-Perry Aggression Questionnaire. *Personality and Individual Differences, 38*, 1631-1643.

DISCUSIÓN

1 Discusión de la adaptación del test BIS-11c

1.1 Sobre la estructura factorial del test BIS-11c

La adaptación colombiana para niños y adolescentes del Bis-11c con 26 ítems realizada a partir de la versión española de Cosi *et al.* (2008b) en general ha tenido un buen ajuste al modelo de tres factores ($\chi^2/gl=2.01$, $RMSEA=.045$, $CFI=.086$), con características similares a la conceptualización teórica inicial de E. S. Barratt (Chahín, Cosi, Lorenzo-Seva, Vigil-Colet, 2010). La razón χ^2/gl ha exhibido un valor inferior a 3.0 que indica un buen ajuste. El índice CFI ha quedado unas cuantas milésimas por debajo del valor de .90 considerado como aceptable. Y el índice $RMSEA$ ha mostrado un valor óptimo, ya que ha estado por debajo de .05 (Hu y Bentler, 1995).

La solución factorial del modelo colombiano replica los factores sobre los cuales Barratt (1985) desarrolló el BIS-10, como son *impulsividad motora*, *impulsividad cognitiva* e *impulsividad no planificadora* (Chahín, 2009). Es de señalar, que la presente versión, así como la española de Cosi *et al.* (2008b), no han replicado la solución bifactorial de Fossati *et al.* (2002) para adolescentes, que plantea una estructura de seis factores de primer orden y dos de segundo orden. Donde el primer factor de segundo orden *impulsividad general*, se halla conformado por los factores de primer orden *impulsividad motora*, *falta de demora en obtener gratificación inmediata*, *atención* y *perseverancia*, que corresponden a los factores *impulsividad motora* e *impulsividad atencional* descritos por Patton *et al.* (1995). Y el segundo

factor de segundo orden, *impulsividad no planificadora*, se encuentra compuesto por los factores de primer orden *complejidad cognitiva* y *autocontrol*, y encajan con la *impulsividad no planificadora* del mismo modelo.

En la presente adaptación colombiana, las escalas incluyeron los mismos ítems que la española. La *impulsividad motora* tuvo ítems de *impulsividad motora*, *atención*, *falta de demora en la gratificación inmediata* y *perseverancia*. La *impulsividad no planificadora* ítems de *complejidad cognitiva* y *autocontrol*. Y en el tercer factor, *impulsividad cognitiva*, los ítems no estuvieron relacionados con ninguno de los factores propuestos para la versión italiana (Chahín, 2009; Chahín *et al.*, 2010; Cosi, 2008; Cosi *et al.*, 2008b; Cosi *et al.*, 2008c; Fossati *et al.*, 2002).

Al aplicar a la versión española y colombiana la rotación Consensus Direct Oblimin (Lorenzo-Seva *et al.*, 2002), se obtuvo una alta congruencia en la mayoría de los ítems, y todos cargaron en los mismos factores en las dos muestras. Las congruencias fueron de .97 para *impulsividad motora*, .90 para *impulsividad cognitiva* y .95 para *impulsividad no planificadora*, lo que indica una alta similaridad entre las estructuras factoriales. Es de aclarar que estas altas congruencias son el resultado del proceso de adaptación que ha implicado cambiar el 80% de los ítems. Dichos cambios fueron realizados teniendo muy presente que las palabras y expresiones utilizadas en la versión española no necesariamente son equiparables a la muestra colombiana, ya que los constructos psicológicos, del mismo modo que sus comportamientos asociados, como en este caso la impulsividad, son dependientes de aspectos culturales y tienen expresiones lingüísticas diferentes (Barbero *et al.*, 2008; Chahín *et al.*, 2010;

Hambleton, 2005; Muñiz y Hambleton, 1996; Van de Vijver y Poortinga, 1997; Van de Vijver y Leung, 2000). Y el resultado, como se ha visto en las congruencias de la rotación Consensus Direct Oblimin, da un buen indicio de que el constructo evaluado es el mismo para las dos culturas. Aspecto que también es ratificado en los indicadores de ajuste resultantes del análisis factorial confirmatorio, como se verá más adelante.

Se ha de aclarar, que aunque la estructura trifactorial original se ha repetido en esta versión como en la española, existen varios aspectos que hay que tener presentes. Primero, se han eliminado varios de los ítems que en su momento fueron elaborados para realizar la adaptación para adolescentes desde la versión italiana para adultos. Segundo, posee dos ítems que no corresponden a ninguna de las dos versiones del instrumento, la original y la italiana para adolescentes, y que fueron redactados sobre la base de una escala de *impulsividad funcional*. Sin embargo, hay que recordar que desde las primeras fases del proceso de adaptación a la población española, y antes de la adición de nuevos ítems, se pudieron identificar las tres dimensiones originales (Barratt, 1985; Brunas-Walgstaff *et al.*, 1997; Chahín *et al.*, 2010; Cosi, 2008, 2011; Cosi *et al.*, 2008b; Cosi *et al.*, 2008c; Fossati *et al.*, 2002).

A pesar de lo anterior, lo que se observa en la versión colombiana es que, también aquí como en la española, las escalas *impulsividad motora* y *no planificadora*, que además presentan una buena correlación positiva entre sí de $r=.35$, se encuentran más relacionadas con aspectos patológicos de la impulsividad. Cuestión que es similar a lo planteado por Eysenck y Eysenck (1977) con la conceptualización

de *impulsividad estricta*, y por Dickman (1990) con *impulsividad disfuncional*. Por otra parte la escala *impulsividad cognitiva* de la presente adaptación, que como se ha dicho coincide con lo informado originalmente por Barratt (1985), implica toma de decisiones rápidas, pero con una característica fundamental, que esta clase de impulsividad es adaptativa, al contrario de las otras dos, como así lo expresan sus cinco ítems (*Decido las cosas rápidamente, Cuando mis amigos me preguntan algo, puedo responder rápidamente, Pienso rápidamente, Me concentro rápidamente, Soy de los primeros en levantar la mano en clase cuando el profesor hace una pregunta*). Este factor cognitivo, como se puede ver en el contenido de los ítems, no se halla asociado a aspectos atencionales considerados negativos, sino a la rapidez en la manera de pensar y de decidir, y se relaciona más con *impulsividad funcional* (Cosi *et al.*, 2008a; Brunas-Wagstaff *et al.*, 1997; Dickman, 1990). Es por esto que las correlaciones de este factor son negativas con la escala *impulsividad motora* ($r=.15$) y con la escala *impulsividad no planificadora* ($r=-.24$) (Chahín *et al.*, 2010). En resumen, las dimensiones *impulsividad motora* e *impulsividad no planificadora* de la adaptación colombiana están asociadas con los aspectos disfuncionales de la impulsividad, mientras que *impulsividad cognitiva* es referida sólo a aspectos funcionales de la impulsividad.

Sobre las razones por la cuales la estructura factorial de la versión colombiana para niños y adolescentes, no fue similar a la propuesta por Patton *et al.* (1995), hay otros aspectos a tener en cuenta. El primero, como se ha descrito en la introducción teórica de este documento, la estructura factorial del BIS-11 no ha sido un constructo consistente, encontrándose soluciones factoriales en las que no siempre se puede

confirmar la estructura trifactorial. En este sentido la escala *impulsividad atencional* ha sido la más problemática en su identificación y la *impulsividad motora* la más consistente entre las muestras (Vasconcelos *et al.*, 2012). Esto puede ser debido a que la solución planteada por Patton *et al.* (1995) de *impulsividad atencional*, *impulsividad motora* e *impulsividad no planificadora*, no es un constructo sólido que se repite en muestras con diferentes características. La situación no cambia significativamente en las investigaciones llevadas a cabo con niños y adolescentes tanto a nivel exploratorio y confirmatorio. Se encuentran algunas investigaciones que reportan un mejor ajuste a los tres factores originales en comparación con otros modelos (Recio, 2012). En otros estudios se identifican sólo dos factores como el de Von Diemen *et al.* (2007). La estructura factorial original no se encontró en la versión italiana para adolescentes, donde fueron hallados seis factores de primer orden y dos de segundo orden, en vez de tres (Fossati *et al.*, 2002). Y situación similar se ha reportado en el análisis factorial confirmatorio realizado en la versión china para adolescentes, en donde también se han reportado seis factores de primer orden y dos de segundo orden (Yao *et al.*, 2007). En resumen, así como ocurre en adultos, los factores del BIS-11 en niños y adolescentes no son estables entre las muestras estudiadas. En consecuencia, este es un aspecto que requiere de más investigación, en especial en lo referente a la diferenciación entre *impulsividad atencional* considerada como *impulsividad disfuncional* e *impulsividad cognitiva* considerada como *impulsividad funcional*, que ha sido uno de los aspectos más problemáticos en el desarrollo la escala.

El segundo aspecto a tener en cuenta referente a las razones por las cuales la estructura factorial obtenida en adultos ha sido diferente de las versiones para niños y adolescentes, tiene que ver con el propio desarrollo evolutivo humano: nos referimos a que la impulsividad no es un rasgo estable durante toda la vida, de hecho cambia con la edad, haciéndose más diferenciada a medida que los niños se hacen mayores (Leshem y Glicksohn, 2007). Por lo tanto los componentes no son fácilmente identificables (Fossati *et al.*, 2002), en especial aquellos que involucran la funcionalidad del rasgo (Brunas-Wagstaff *et al.*, 1997).

1.2 Aspectos concernientes a las comparaciones multigrupo

Uno de los temas de trabajo de esta investigación ha sido el comprobar si las dos versiones del instrumento (la española y la colombiana) miden el mismo constructo de impulsividad. Una vez obtenido el modelo factorial para cada muestra y comprobada la respectiva estructura trifactorial, se llevaron a cabo una serie de pasos jerárquicos secuenciales, en los que en cada uno se agregaban mayores restricciones para determinar la equivalencia existente entre los dos soluciones factoriales (Byrne, 2008; Chahín *et al.*, 2010; Cosi *et al.*, 2008b; Cosi *et al.*, 2008c; Millsap y Meredith, 2007).

En principio, con la invarianza configuracional (*configurational invariance*) se ha comprobado la equivalencia de los parámetros estructurales, es decir la condición invariante en el patrón de la matriz de tres factores del BIS-11c entre la muestra española y la colombiana. Consiguiéndose un muy buen indicador de ajuste absoluto

en la raíz cuadrada de la media del error de aproximación ($RMSEA=.045$), ya que su valor ha sido inferior a .05. El índice comparativo de ajuste estuvo dos centésimas por debajo del punto de corte de .9 con un valor $CFI=.88$. El χ^2/gl fue de 2.05, valor inferior a tres, que informa igualmente un buen ajuste. Estos resultados indican que los tres factores del BIS-11c son los mismos para los dos muestras analizadas (Bentler, 1990; Browne, y Cudek, 1993; Chahín *et al.*, 2010; Fan y Sivo, 2007; Hu y Bentler, 1999).

Se procedió posteriormente a utilizar un método más restrictivo, como es la prueba de invarianza factorial débil (*weak factorial invariance*), que involucra principalmente variables observadas y su relación con las variables latentes, e indica que cualquier diferencia sistemática en las covarianzas entre las medidas de las variables es debida a factores comunes, más que a otras fuentes de asociación. El valor de la raíz cuadrada de la media del error de aproximación ($RMSEA=.049$) continúa dentro de los límites habitualmente aceptados como buenos. No obstante, no ocurre lo mismo con el índice comparativo de ajuste ($CFI=.85$) que está algo más por debajo de lo propuesto para el punto de corte. A pesar de lo expuesto, no es pertinente rechazar la invarianza factorial débil, sólo por haberse obtenido un CFI inferior a .90, ya que no existe unanimidad respecto a cuáles son los valores del CFI que pueden ser tomados como adecuados. Los obtenidos en esta investigación son considerados por Little (1997) como indicadores de un buen ajuste, utilizando para llegar a esta conclusión, la comparación de los CFI de la invarianza factorial débil con la invarianza configuracional. Complementando el procedimiento seguido, se tuvo en cuenta un acercamiento desde otra perspectiva, que consintió en incorporar análisis

provenientes del análisis factorial exploratorio, como es el caso la rotación Consensus Direct Oblimin, que permite determinar el grado de congruencia de las dos soluciones factoriales (Lorenzo-Seva *et al.*, 2002). Los valores del grado congruencia encontrados entre los factores han informado una buena similitud para todos los factores en conjunto (.94) y para cada uno de ellos específicamente (*impulsividad motora* .97, *impulsividad no planificadora* .95 e *impulsividad cognitiva* .90). Utilizando esta técnica exploratoria, igualmente se concluye que los tres factores del BIS-11c son equivalentes en las dos muestras. En consecuencia, aunque el *CFI* se encuentra por debajo del punto de corte, se puede considerar que los factores son equivalentes entre las dos versiones, y se podría concluir bajo ciertas restricciones que las cargas de los factores son proporcionalmente iguales en los dos grupos (Bentler, 1990; Chahín *et al.*, 2010; Fan y Sivo, 2007; Hu y Bentler, 1999; Little, 1997).

Y finalmente, al realizar la comparación con una invarianza factorial fuerte (*strong factorial invariance*), se ha intentado confirmar si las cargas factoriales y los interceptos son proporcionalmente iguales entre los dos grupos. En este caso aunque el indicador de ajuste absoluto continúa siendo aceptable (*RMSEA*=.061), no ocurre lo mismo con el ajuste relativo que se encuentra ya muy distante del punto de corte aceptable (*CFI*=.73). Por lo tanto, las diferencias entre los puntajes obtenidos de las dos versiones del BIS-11c, no pueden ser atribuidas a diferencias reales del constructo impulsividad (Bentler, 1990; Browne y Cudek, 1993; Van de Vijver y Tanzer, 2004). El pobre indicador puede ser causado en este caso, por sesgos de método que no permitieron que el intercepto fuera el mismo para las dos culturas. Creemos que uno de los sesgos que ha podido tener un mayor peso ha sido el que las muestras no fueron

homogéneas en sus características: específicamente la edad, donde las diferencias son bastante notorias. Aquí las edades de los niños españoles se encuentran entre los 9 y 13 años ($M=11.0$, $DT=.92$), mientras que en los colombianos el rango es más amplio, entre 8 y 16 años ($M=13.0$, $DT=2.38$). Una diferencia en las medias de dos años en estas edades genera un alto sesgo, ya que las características del rasgo impulsividad varían con la edad. Como se ha descrito, diversos autores señalan que el constructo es menos diferenciado en la niñez y la adolescencia, mejorando a medida que la edad incrementa, convirtiéndose en un rasgo consistente en la adultez (Brunas-Walgstaff *et al.*, 1997; Fossati *et al.*, 2002; Leshem y Glicksohn 2007). Por consiguiente, los interceptos y las cargas factoriales no pueden ser los mismos, si las edades de las muestras tienen diferencias tan marcadas (Chahín *et al.*, 2010).

Es de aclarar que el objetivo más importante de la presente investigación no fue realizar una comparación transcultural del BIS-11c, sino desarrollar una versión adaptada a las características propias de la población de niños y adolescentes colombianos, a partir de la versión propuesta para la correspondiente población española. Si en el futuro se planteara un nuevo estudio que busque realizar comparaciones entre los puntajes obtenidos entre la población española y la colombiana, se podrá utilizar la presente versión colombiana, que ha tenido una adaptación cultural rigurosa, pero al mismo tiempo deberán tenerse en cuenta los aspectos metodológicos, para que se cumplan todos los criterios de equivalencia multigrupo en las muestras comparadas.

1.3 Consistencia interna del test BIS-11c

1.3.1 Las escalas impulsividad motora e impulsividad no planificadora

Esta adaptación ha tenido una consistencia interna aceptable en la escala *impulsividad motora* con $\alpha=.74$ (intervalo de confianza entre .70 y .77) e *impulsividad no planificadora* con $\alpha=.72$ (intervalo de confianza entre .68 y .75) (Chahín, 2009; Chahín *et al.*, 2010). Y aunque los ítems y los factores no corresponden exactamente a otras versiones existentes, y ante la falta de estudios disponibles que reporten la invarianza del modelo entre las distintas versiones, puede ser aclaratorio realizar algunas comparaciones con lo revisado aquí para cada una de las escalas, con el objeto de poder tener algún punto de contrastación respecto a la consistencia interna obtenida en la adaptación colombiana.

Para empezar, al confrontar la consistencia de la *impulsividad motora* de la versión colombiana ($\alpha=.74$), con la obtenida en las muestras adultas citadas, en las que se menciona específicamente esta escala, se ve que su valor se encuentra por encima de lo reportado en investigaciones como la norteamericana de Stanford *et al.* (2009) con $\alpha=.59$, china de Hou *et al.* (2006) con $\alpha=.65$, japonesa de Someya *et al.* (2001) con $\alpha=.64$, argentina de Folino *et al.* (2006) con $\alpha=.70$, turca de Güleç *et al.* (2008) con $\alpha=.38$ y $\alpha=.27$ y española de Luengo *et al.* (1991) con $\alpha=.63$ y $\alpha=.71$. Por otra parte, al compararla con muestras de niños, hay dos estudios, el primero, que

informó un valor inferior, como es el caso de Recio (2012) con $\alpha=.67$, y el segundo, el de la versión base de Cosi *et al.* (2008b) con un coeficiente superior de $\alpha=.80$. En lo relacionado con la fiabilidad de la escala *impulsividad no planificadora* ($\alpha=.72$), las mismas publicaciones señalan unos valores iguales o menores en adultos, como Stanford *et al.* (2009) con $\alpha=.72$, Hou *et al.* (2006) con $\alpha=.68$, Someya *et al.* (2001) con $\alpha=.65$. Folino (2006) con $\alpha=.48$, Güleç *et al.* (2008) con $\alpha=.23$ y $\alpha=.43$ y Luengo *et al.* (1991) con $\alpha=.20$ y $\alpha=.10$. Y en lo referente a los niños, sus indicadores son superiores a los dos estudios españoles, el de Recio (2012) con $\alpha=.63$ y el de Cosi *et al.* (2008b) con $\alpha=.71$. Con lo expuesto en este párrafo se puede ver que estas dos escalas de la versión colombiana, *impulsividad motora* e *impulsividad no planificadora*, además de presentar unos buenos coeficientes de consistencia interna, sus valores son comparativamente superiores a los revisados, tanto en adultos como en niños. Comprobándose, al menos para estas dos escalas, que en instrumentos autoaplicados en niños (como es el caso del BIS-11c) es posible obtener buenos indicadores de confiabilidad (Cosi *et al.*, 2008a; Fink y McCown, 1993; Zaparniuk y Taylor, 1997). Estos resultados permiten concluir que estas dos dimensiones del BIS-11c, *impulsividad motora* e *impulsividad no planificadora*, poseen un constructo bastante sólido y, dadas sus cualidades psicométricas, pueden ser utilizadas con fines de investigación o clínicos en muestras con características similares en el contexto colombiano.

1.3.2 La escala impulsividad cognitiva

Contrariamente a lo que sucede con las dos escalas anteriores, *impulsividad cognitiva* posee un coeficiente de $\alpha=.59$ (intervalo de confianza entre .55 y .62), que habitualmente se asocia con una consistencia interna pobre. Puede observarse que sólo el valor del límite superior del intervalo de confianza es marginalmente aceptable (Chahín, 2009; Chahín *et al.*, 2010). Hay que recordar que se observó una situación similar en la primera fase del proceso de adaptación de la versión española llevado a cabo a partir de la italiana. La escala obtuvo valores bajos para toda la muestra ($\alpha=.60$) y bastante menores para los grupos de edad de 10 y 11-12 años ($\alpha=.58$ y $\alpha=.56$) (Cosi *et al.*, 2008c). Para mejorar la fiabilidad los autores agregaron dos nuevos ítems a la escala *impulsividad cognitiva*, mejorándola hasta $\alpha=.68$ (Cosi *et al.*, 2008b). Cabe anotar que el problema de la baja fiabilidad de la escala no es exclusivo del desarrollo y adaptación de la versión del BIS-11c para la población de adolescentes españoles o colombianos (Cosi, 2011; Vasconcelos *et al.*, 2012). En otro instrumento como el *Test de Impulsividad* de Dickman, la *impulsividad funcional*, relacionada con la tendencia a tomar decisiones rápidas en situaciones apropiadas, que es similar a la escala *impulsividad cognitiva* del BIS-11c, ha tenido coeficientes de consistencia interna bajos en niños entre 8 y 9 años con $\alpha=.38$, 11 y 12 años con $\alpha=.50$ y 15 y 16 años con $\alpha=.51$ (Brunas-Wagstaff *et al.*, 1997).

La *impulsividad cognitiva* de la presente adaptación, como se ha descrito, ha exhibido una situación similar a la ocurrida inicialmente con la española. Pero hay que tener presente que este problema también se dio en el BIS-10 con adultos cuando

Luengo *et al.* (1991) informaron en sus dos muestras fiabilidades demasiado bajas ($\alpha=.27$ y $\alpha=.18$), al mismo tiempo sus ítems se distribuían entre todos los factores, por lo que se concluyó que dicha escala no podía ser considerada como un factor independiente. Estos autores afirman que *impulsividad cognitiva* está más relacionada con el factor denominado *decisiones rápidas*. Este factor se compone por ítems del BIS-8, del BIS-10 y del I5 (*I answer quickly, I make-up mi mind quickly y Do you usually make up your mind quickly?*). Como se puede observar, son ítems muy similares a los de la versión del BIS-11c, donde el factor cognitivo se encuentra relacionado con tomar decisiones rápidas, en vez de aspectos que implican dificultades en la atención o pensamientos intrusivos (Cosi, 2011; Gerbing *et al.*, 1987).

La baja consistencia interna observada en la escala *impulsividad cognitiva* puede tener varias posibles explicaciones. Barratt (1994) sostiene que las funciones cognitivas son difíciles de medir en instrumentos autoadministrados, mucho más si lo que se evalúa son procesos cognitivos inferenciales. Igualmente, los componentes de la impulsividad cambian con la edad y son más inespecíficos en edades tempranas (Fossati *et al.*, 2002; Leshem y Glicksohn, 2007). Por lo demás, este factor cognitivo no siempre ha sido identificado tal y como fue conceptualizado originalmente (Barratt, 1985; Stanford *et al.*, 2009; Patton *et al.*, 1995; Vasconcelos *et al.*, 2012).

Para mejorar la consistencia interna de la escala *impulsividad cognitiva* en la versión colombiana, y obtener al menos una confiabilidad aceptable que se acerque a la española, es necesario agregar nuevos ítems a los ya existentes, ya que de hecho su

número actual continúa siendo bajo en comparación con las otras dos escalas. Una fuente posible de nuevos ítems son los contenidos en la escala *impulsividad cognitiva* original propuesta por Barratt (1985) para el BIS-10. Y sobre ellos realizar una cuidadosa adaptación para niños y adolescentes, siguiendo las directrices propuestas por Hambleton (2005). Por otra parte, como la *impulsividad cognitiva* tiene una conceptualización teórica similar a la *impulsividad funcional*, se podrían volver a realizar las adaptaciones de los ítems provenientes de dicha escala (Dickmans, 1990; Brunas-Wagstaff *et al.*, 1997), como se hizo en la versión española (Cosi, 2008; Cosi *et al.*, 2008b).

2 Discusión de la adaptación del test AQ

2.1 Sobre la estructura factorial del test AQ

Como se ha descrito, los índices de ajuste del modelo han estado dentro de los parámetros aceptables. En particular, el indicador de ajuste absoluto (el error medio cuadrático de aproximación) ha tenido un valor que revela un muy buen ajuste, ya que se encuentra por debajo límite de .05 ($RMSEA=.049$). Con respecto al indicador de ajuste relativo (el índice de ajuste comparativo) recomendado para medir variables psicológicas, ha arrojado un valor superior a .90 ($CFI=.92$) que informa de un buen ajuste. Del mismo modo, el índice de ajuste normativo, medida de discrepancia entre el modelo ajustado y el modelo base, ha obtenido un valor de $NFI=.90$. Visto lo anterior, puede afirmarse que esta adaptación colombiana de 20 ítems del AQ para

preadolescentes y adolescentes, exhibe más que un buen ajuste al modelo de cuatro factores propuesto por Buss y Perry (1992) para evaluar la agresividad (Bentler, 1990; Bentler y Bonett, 1980; Chahín-Pinzón, Lorenzo-Seva y Vigil-Colet, 2012).

Es pertinente para analizar estos resultados, en primer término comparar el ajuste obtenido aquí con la versión que originó esta adaptación, que tuvo como una de sus características principales estar libre de sesgos debidos a las diferencias culturales. Aunque la adaptación base en Morales-Vives *et al.* (2005) ha mostrado un buen ajuste, llama la atención que sus indicadores, tanto absolutos *RMSEA* (.056 vs .049) como relativos *CFI* (.90 vs .92), han sido comparativamente inferiores a la adaptación colombiana (Chahín-Pinzón *et al.*, 2012). Es decir, que en este caso, el ajuste al modelo de cuatro factores con 20 ítems ha sido ligeramente superior en la muestra de niños y adolescentes colombianos, en comparación a la original realizada con adultos españoles.

El ajuste al modelo obtenido por Chahín-Pinzón *et al.* (2012) ha sido igualmente superior al reportado en adultos con la versión original completa de 29 ítems de Buss y Perry (1992). Son los estudios publicados por Harris (1995) (*GFI*=.825, *AGFI*=.794, *RMSR*=.077), Meesters *et al.* (1996) (*GFI*=.87, *AGFI*=.84, *RMSR*=.064), Williams *et al.* (1996) (*NFI*=.73, *CFI*=.84), Bryant y Smith (2001) (*GFI* entre .76 y .81, *RMSEA* entre .072 y .084, *CFI* entre .81 .82, *NNFI* entre .76 y .80) y Morales-Vives *et al.* (2005) (*RMSEA*=.055, *CFI*=.88, *GFI*=.84, *AGFI*=.82). Por lo demás, la presente adaptación ha exhibido unos indicadores bastante similares a la española de Andreu *et al.* (2002) con *GFI*=.93, *AGFI*=.92 y *RMSEA*=.05, y a los de

Meesters *et al.* (1996) tras eliminar tres de los ítems de la versión original ($GFI=.90$, $AGFI=.88$).

Situación algo diferente se da al revisar los ajustes en la versión refinada de 12 ítems y en algunas de sus posteriores adaptaciones en población adulta. Éste es el caso de Bryant y Smith (2001), en donde los indicadores $RMSEA$ expresaron en cuatro de la muestras seleccionadas un ajuste inferior, que estuvo entre .063 y .094, mientras los indicadores de ajuste relativo fueron ligeramente superiores para algunos casos con CFI entre .91 y .96 y $NNFI$ entre .87 y .94. Del mismo modo, Vitoratou *et al.* (2009) encontraron ajustes inferiores con GFI entre .71 y .86, $RMSEA$ entre .067 y .096, $NNFI$ entre .57 y .76 y CFI entre .61 y .75. Es de resaltar que en española de Gallardo-Pujol *et al.* (2006) los indicadores mejoraron ($RMSEA=.053$, $CFI=.93$), siendo cercanos a los obtenidos para niños colombianos por Chahín-Pinzón *et al.* (2012).

Ahora, al cotejar los índices de esta adaptación con versiones largas que han trabajado con adolescentes y preadolescentes, los resultados son contradictorios (Chahín *et al.*, 2012). Algunos de ellos reportan un ajuste muy inferior al obtenido en la población colombiana, como Hornsveld *et al.* (2009) con $GFI=.72$, $CFI=.65$ y $RMSEA=.09$. Otros dan cuenta de un ajuste cercano, como son en lengua española el estudio de Santisteban y Alvarado (2009) con $GFI=.92$, $AGFI=.91$ y $RMSEA=.047$ y el de Reyna *et al.* (2011) con $GFI=.90$, $AGFI=.88$, $CFI=.97$ y $RMSEA=.042$.

Y respecto a las versiones refinadas del AQ para niños y adolescentes, Hornsveld *et al.* (2009) notificaron un ajuste menor con $GFI=.93$, $CFI=.94$ y $RMSEA=.06$. Llama la atención que la versión en idioma chino de Ang (2007) ha exhibido ajustes superiores en sus dos muestras con CFI entre $.96$ y $.98$ y $RMSEA$ entre $.047$ y $.37$. Lo mismo ocurrió con la muestra argentina de Reyna *et al.* (2011) con $CFI=.99$ y $RMSEA=.027$.

Creemos que de acuerdo a lo discutido hasta ahora, el resultado de un superior ajuste al modelo de cuatro factores de Buss y Perry (1992) de la versión colombiana puede ser por varias razones. La primera, debida al proceso llevado a cabo para la adaptación de los ítems, en los que se tuvo que realizar una muy exigente adaptación cultural, que involucró expertos en psicometría, conocedores de las dos culturas, junto con un grupo de profesores con amplia experiencia y conocimiento del vocabulario y desarrollo evolutivo de los niños. Este complejo proceso implicó entre otras cosas, cambiar palabras, expresiones y frases, de tal forma que se pudiera garantizar que el sentido y significado de cada uno de los reactivos reflejara de la manera más precisa la dimensión a la que pertenecía, de tal forma que se controlara este tipo de sesgos. Complementando lo anterior, se tuvo especial cuidado en adaptar adecuadamente los ítems para las características de la población preadolescente y adolescente, ya que la versión española fue desarrollada para adultos, por lo que se tuvo en cuenta que los ítems estuvieran contextualizados a la características propias de esas edades. La segunda razón, es que esta adaptación de 20 ítems está libre de los dos ítems invertidos que están presentes en la original, que han demostrado reiterativamente que afectan el ajuste del modelo. Y la tercera razón, que involucra en algo la anterior,

tiene que ver con que la versión española ha sido desarrollada teniendo en cuenta sólo los ítems libres de sesgos culturales, con lo que se garantiza que estos funcionen adecuadamente en otras culturas, como ha ocurrido con la adaptación para Colombia.

2.2 Consistencia interna del test AQ

2.2.1 Comparación de la consistencia interna con las versiones para adultos

Es conveniente primero, para esta parte de la discusión, comparar los indicadores de la consistencia interna de Chahín-Pinzón *et al.* (2012) con los dados por Morales-Vives *et al.* (2005). Advertimos que los coeficientes para la escala total en las dos versiones han sido adecuados, siendo superior para la española ($\alpha=.82$ vs $\alpha=.87$). Igualmente las fiabilidad de las escalas han sido superiores: *agresividad física* ($\alpha=.75$ vs $\alpha=.88$), *agresividad verbal* ($\alpha=.58$ vs $\alpha=.71$), *ira* ($\alpha=.58$ vs $\alpha=.68$) y *hostilidad* ($\alpha=.62$ vs $\alpha=.65$). Indudablemente como se ve, los coeficientes obtenidos en la muestra de adultos españoles son mejores. No obstante, no hay que perder de vista que se están comparando muestras con medias de edad muy diferentes: una edad media de 27.6 ($DT=1.14$) en la muestra española, respecto a una edad media de 12.7 ($DT=2.11$) en la muestra colombiana. Valores inferiores en el coeficiente de Cronbach suelen darse cuando un instrumento (en este caso el AQ) ha sido adaptado a poblaciones de niños y adolescentes, aspecto que se discutirá más adelante. Pero, si únicamente se analiza el segmento de la muestra entre 14 y 16 años, la cuestión

mejora, exhibiendo valores bastante cercanos a los adultos, tanto para la escala total ($\alpha=.83$ vs $\alpha=.87$) como para las respectivas escalas: *agresividad física* de ($\alpha=.77$ vs $\alpha=.88$), *agresividad verbal* de ($\alpha=.66$ vs $\alpha=.71$), *ira* ($\alpha=.63$ vs $\alpha=.68$) y *hostilidad* ($\alpha=.65$ vs $\alpha=.65$).

A continuación se confrontarán los resultados de la investigación actual publicada en Chahín-Pinzón *et al.* (2012) con diversos estudios que se han efectuado con adultos.

En primera instancia, se van a comparar los resultados referidos a la escala total del AQ. En este aspecto la versión colombiana ostenta una buena fiabilidad con $\alpha=.82$, y el coeficiente es consistente en otras muestras adultas. Cabe apuntar que el valor es inferior a los obtenidos por Morales-Vives *et al.* (2005) con $\alpha=.87$ y Buss y Perry (1992) con $\alpha=.89$. Por otra parte, es igual a los de García-León *et al.* (2002), Castrillón *et al.* (2005) y Sommantico *et al.* (2008) con $\alpha=.82$. Y finalmente, es superior entre otras a las revisiones de Bryant y Smith (2001) en español comunicadas por Gallardo-Pujol *et al.* (2006) con $\alpha=.78$ y por Morales-Vives *et al.* (2005) con $\alpha=.81$.

En segunda instancia, se examinará la escala *agresividad física*. Ésta ha obtenido un indicador adecuado con $\alpha=.75$ en la muestra colombiana, pero inferior al de Morales-Vives *et al.* (2005) con $\alpha=.88$ y al de Buss y Perry (1992) con $\alpha=.82$. Por otra parte, es superior a las vistas en versiones no refinadas como las de Porras *et al.* (2001) con $\alpha=.70$, García-León *et al.* (2002) con $\alpha=.63$, Harris (1995) con $\alpha=.75$,

Meesters *et al.* (1996) con $\alpha=.75$, Castrillón *et al.* (2005) con $\alpha=.81$ y Vitoratou *et al.* (2009) con $\alpha=.82$. Al mismo tiempo ha sido superior en las versiones reducidas de Bryant y Smith (2001) con $\alpha=.79$ y $\alpha=.80$, Vitoratou *et al.* (2009) en su tres muestras con $\alpha=.77$, $\alpha=.70$ y $\alpha=.76$ y a la española de Gallardo-Pujol *et al.* (2006) con $\alpha=.70$.

En tercera instancia, la consistencia interna de la escala *agresividad verbal* de la presente versión es baja con $\alpha=.58$, e inferior a la obtenida por Morales-Vives *et al.* (2005) con $\alpha=.71$ y por Buss y Perry (1992) con $\alpha=.72$. Al compararla con otros estudios, se advierte que este valor bajo no es exclusivo de esta adaptación, sino que también se da en varias publicaciones con adultos. Por ejemplo, en lo que se refiere a las versiones no refinadas, están la de Fossati *et al.* (2003) con $\alpha=.53$, Meesters *et al.* (1996) con $\alpha=.50$ y Vitoratou *et al.* (2009) en sus tres muestras con $\alpha=.55$, $\alpha=.50$ y $\alpha=.51$. Lo mismo ocurre en las refinadas con 12 ítems, donde las cosas no fueron diferentes. Gallardo-Pujol *et al.* (2006) informaron un coeficiente de $\alpha=.58$ y Vitoratou *et al.* (2009) de $\alpha=.55$, $\alpha=.56$ y $\alpha=.48$. No obstante, si tomamos el valor de la fiabilidad del grupo de mayor edad (entre 14 a 16 años) comunicado por Chahín-Pinzón *et al.* (2012) tenemos un $\alpha=.66$. Y este valor supera los coeficientes de las versiones largas de García-León *et al.* (2002) con $\alpha=.63$ y Porras *et al.* (2001) con $\alpha=.62$, aproximándose al obtenido por Andreu *et al.* (2002) con $\alpha=.68$. Igualmente en la evaluación de la versión refinada de 12 ítems de Morales-Vives *et al.* (2005) el coeficiente informado ($\alpha=.63$) se sitúa por debajo del valor obtenido en este grupo de edad.

En cuarta instancia, el coeficiente de la escala *ira* en esta adaptación ha sido bajo con $\alpha=.58$, e inferior al de Morales-Vives *et al.* (2005) con $\alpha=.68$ y al de Buss y Perry (1992) con $\alpha=.83$. Más hay que agregar, que el valor obtenido aquí no es consistente con las investigaciones de esta escala en adultos. Al intentar comparar este coeficiente se aprecia que todas las publicaciones revisadas con adultos tienen fiabilidades más altas. La excepción es la de Vitoratou *et al.* (2009) con valores inferiores ($\alpha=.57$ y $\alpha=.54$). Incluso el valor para el segmento de mayor edad (entre 14 y 16 años) que es de $\alpha=.63$, continúa siendo inferior en todos los estudios analizados, en versiones refinadas y en largas, aunque acercándose ligeramente a los coeficientes de las revisiones de Bryant y Smith (2001) llevadas a cabo por Gallardo-Pujol *et al.* (2006) con $\alpha=.66$ y por Morales-Vives *et al.* (2005) con $\alpha=.66$.

En quinta y última instancia, en lo que se refiere a la escala *hostilidad*, ésta ha obtenido un coeficiente bajo de $\alpha=.62$, que es inferior al de Morales-Vives *et al.* (2005) con $\alpha=.65$ y al de Buss y Perry (1992) con $\alpha=.77$. En esta escala el valor es menor a los informes vistos, tanto en versiones largas como reducidas, con excepción nuevamente de Vitoratou *et al.* (2009) con $\alpha=.61$. La situación continúa siendo similar en el segmento de edad entre 14 y 16 años ($\alpha=.65$), que sigue estando por debajo de los estudios examinados, y sólo se iguala a la de Morales-Vives *et al.* (2005), y se acerca un poco a la de Gallardo-Pujol *et al.* (2006) con $\alpha=.68$.

Lo que se puede concluir hasta este momento sobre los coeficientes alfa Cronbach de la presente adaptación para niños y adolescentes del AQ de Chahín-Pinzón *et al.* (2012), es que al compararla con la información disponible de muestras

con adultos, la escala total y la escala *agresividad física* y en cierta medida la escala *agresividad verbal*, son consistentes con las investigaciones disponibles en la literatura. En cambio, la escala *ira* y la escala *hostilidad*, poseen en general valores substancialmente más bajos que los observados en los adultos. Los coeficientes mejoran, aunque no lo suficiente, cuando sólo se tiene en cuenta para el análisis el rango de mayor edad de la muestra colombiana, y por el contrario, empeoran, cuando la comparación es realizada con los otros dos grupos de edad menor.

2.2.2 Comparación de la consistencia interna con las versiones para niños y adolescentes

A partir de este punto se desarrollará el análisis y la discusión de la fiabilidad de la versión de Chahín-Pinzón *et al.* (2012), contrastándola con las existentes del AQ que se han realizado para niños y adolescentes. Al hacer la comparación de la consistencia de escala total ($\alpha=.82$) lo que se encuentra son diferencias mucho más favorable de las que se vieron anteriormente con adultos. Es conveniente remarcar aquí, que aunque esta adaptación colombiana no ha superado las consistencias de las versiones largas, sí que ha logrado coeficientes iguales o superiores a los observados en casi todas las versiones refinadas para estas edades: como son las de Hornsveld *et al.* (2009) con $\alpha=.81$, Ang (2007) en una de sus dos muestras con $\alpha=.82$ y en español Reyna *et al.* (2011) con $\alpha=.82$.

Se considera importante realizar una discusión un poco más exhaustiva de lo informado por Chahín-Pinzón *et al.* (2012) en lo que se refiere a cada una de las escalas.

Primero, la escala *agresividad física* que en esta versión ha obtenido un $\alpha=.75$, y en los tres grupos de edad continúa reportando una consistencia interna, se encuentra dentro de unos parámetros bastante adecuados. Incluso en edades muy tempranas (8 y 10 años con $\alpha=.80$, 11 y 13 años con $\alpha=.73$, 14 y 16 años con $\alpha=.77$), aspecto que de hecho es poco común. Al comparar estas fiabilidades con otras muestras de niños y adolescentes, las diferencias cambian de acuerdo a la cantidad de ítems de la versión con la que se equipara. En lo que implica a las reducidas, como son las adaptaciones para adolescentes hechas a partir de Bryant y Smith (2001), los coeficientes para los tres grupos de edad y para el total de la adaptación colombiana son superiores en todos los casos, como se aprecia en las dos muestras de Ang (2007) quien informó confiabilidades de $\alpha=.56$ y $\alpha=.72$, Hornsveld *et al.* (2009) con $\alpha=.62$ y en la argentina de Reyna *et al.* (2011) con $\alpha=.70$. Y desde el punto de vista de las que han utilizado todos o gran parte de los ítems del AQ, la presente adaptación ha logrado valores iguales o superiores a los de Morren y Meesters (2002) con $\alpha=.75$ y Santisteban *et al.* (2007) en una de las muestras de mujeres preadolescentes con $\alpha=.70$. Pero por otro lado en versiones largas, la colombiana ha arrojado puntajes inferiores a los obtenidos por Hornsveld *et al.* (2009) con $\alpha=.78$, López del Pino *et al.* (2009) con $\alpha=.85$, Fossati *et al.* (2003) con $\alpha=.81$ y Reyna *et al.* (2011) con $\alpha=.77$. De igual modo, Santisteban *et al.* (2007) en sus otras tres muestras informaron unos coeficientes superiores para adolescentes hombres con $\alpha=.83$ y mujeres $\alpha=.82$, y para

preadolescentes hombres con $\alpha=.78$. Finalmente, de acuerdo a lo expuesto, consideramos que la escala *agresividad física* funciona bastante bien en la adaptación colombiana, tanto para adolescentes como preadolescentes, ya que aparte de poseer una buena fiabilidad, exhibe mejores indicadores que las otras versiones refinadas, e incluso en algunos casos sus valores son iguales o superiores a las versiones no reducidas.

Segundo, la escala *agresividad verbal* en la versión colombiana ha obtenido una consistencia interna baja con $\alpha=.58$, y para los grupos de menor edad el valor ha sido inferior, de 8 y 10 años con $\alpha=.48$ y de 11 y 13 años con $\alpha=.51$, mientras que en los de 14-16 años ha mostrado un mejor comportamiento con $\alpha=.66$. Al comparar los resultados, éstos son bastante contradictorios entre los diversos artículos. Con las versiones refinadas advertimos que una de las muestras chinas de Ang (2007) ha alcanzado un valor mucho más bajo con $\alpha=.50$. Lo mismo le ha ocurrido a Reyna *et al.* (2011) con $\alpha=.55$. Inversamente, Hornsveld *et al.* (2009) han hallado un valor levemente superior con $\alpha=.59$, mientras que la otra muestra china de Ang (2007) reporta un valor más alto con $\alpha=.70$. En lo que se refiere a la comparación con las versiones largas, éstas han presentado menores coeficientes: Morren y Meesters (2002) con $\alpha=.51$, Fossati *et al.* (2003) con $\alpha=.53$ y Hornsveld *et al.* (2009) con $\alpha=.34$. Por otro lado han logrado mayores valores: López del Pino *et al.* (2009) con $\alpha=.64$, Reyna *et al.* (2011) con $\alpha=.68$ y Santisteban *et al.* (2007) con preadolescentes (hombres con $\alpha=.68$ y mujeres con $\alpha=.66$) y adolescentes (hombres con $\alpha=.76$ y mujeres con $\alpha=.72$). Pero hay que tener presente, que si se toma como base de la comparación el grupo de mayor edad de 14 y 16 años con $\alpha=.66$, la cuestión cambia y

los valores son muy próximos a la muestra de preadolescentes españoles de Santisteban *et al.* (2007). Finalmente, aunque la consistencia interna de la escala *agresividad verbal* para toda la muestra ha sido baja, esto no es exclusivo sólo de la colombiana: este tipo valores es frecuente en estudios para adolescentes y preadolescentes tanto en versiones cortas como largas. Por lo que se puede considerar que el valor obtenido se enmarca dentro de lo esperado para *agresividad verbal* en estas edades. No obstante, cabe resaltar que el coeficiente ha sido superior a lo reportado en diversos informes con población similar. Además, la situación tiende a mejorar cuando se toma como referencia el grupo de mayor edad, ya que se observa una mejoría de la consistencia a medida que incrementa la edad de los niños, y va cambiando de valores inaceptables (en los rangos de menor edad) hasta marginalmente aceptables (en los rangos de mayor edad).

Tercero, la escala *ira* también ha presentado una consistencia interna baja con $\alpha=.58$. En los grupos de menor edad tuvo valores igualmente inaceptables para los grupos de 8 y 10 años con $\alpha=.36$ y 11 y 13 años con $\alpha=.58$, logrando una fiabilidad marginalmente aceptable en el grupo de 14 y 16 años con $\alpha=.63$. Resulta interesante que, a pesar de esta baja consistencia, al comparar la escala *ira* con las otras versiones reducidas, la colombiana presenta coeficientes superiores en todos los artículos que se han estudiado: es el caso de las dos muestras de Ang (2007) con $\alpha=.34$ y $\alpha=.55$, Hornsveld *et al.* (2009) con $\alpha=.51$ y Reyna *et al.* (2011) con $\alpha=.48$. El panorama es muy distinto en las versiones largas, en donde todas ellas ostentan valores superiores a los obtenidos en la presente adaptación: Morren y Meesters (2002) con $\alpha=.67$, Fossati *et al.* (2003) con $\alpha=.72$, Hornsveld *et al.* (2009) con $\alpha=.64$, López del Pino *et al.*

(2009) con $\alpha=.78$, Reyna *et al.* (2011) con $\alpha=.67$ y Santisteban *et al.* (2007) en preadolescentes (hombres con $\alpha=.61$ y mujeres con $\alpha=.61$) y adolescentes (hombres con $\alpha=.66$ y mujeres con $\alpha=.69$). Si bien es importante mencionar que la fiabilidad en el rango de edad entre 14 y 16 años de la muestra colombiana ($\alpha=.63$) es superior a los valores de las dos muestras de preadolescentes de Santisteban *et al.* (2007). Finalmente, aunque la consistencia interna encontrada aquí para la escala *ira* no ha sido adecuada, es de puntualizar que la escala ha presentado desde un comienzo valores bajos en adultos, y se hace más evidente cuando se trabaja con niños y adolescentes. El problema se acentúa aun más en las versiones reducidas, donde la consistencia interna es usualmente menor. Por tanto, lo hallado en la adaptación colombiana es consistente con otras investigaciones en este rango de edades. No obstante lo anterior, hay que puntualizar que la presente versión ha alcanzado puntajes superiores a todas las versiones reducidas analizadas. Y en el rango de mayor edad, incluso los coeficientes son superiores a algunas de las versiones largas, lo que indica como el comportamiento de esta escala mejora con la edad, desde la niñez hasta la adolescencia, empezando sólo a ser mínimamente aceptable a partir de los 14 años.

La escala *hostilidad* ha mostrado un coeficiente $\alpha=.62$, con valores que continúan siendo aceptables para los rangos de edad de 8 y 10 años con $\alpha=.61$, 11 y 13 años con $\alpha=.60$ y 14 y 16 años con $\alpha=.65$. Con respecto a las versiones refinadas, esta escala tiene puntajes superiores a las dos muestras de Ang (2007) con $\alpha=.38$ y con $\alpha=.61$ y Reyna *et al.* (2011) con $\alpha=.54$, y posee un valor inferior al obtenido por Hornsveld *et al.* (2009) con $\alpha=.74$. En las versiones largas exhibe coeficientes inferiores a los de Morren y Meesters (2002) con $\alpha=.69$, Fossati *et al.* (2003) con

$\alpha=.68$, Hornsveld *et al.* (2009) con $\alpha=.73$, López del Pino *et al.* (2009) con $\alpha=.78$ y Reyna *et al.* (2011) con $\alpha=.71$. Santisteban *et al.* (2007) de igual forma informan de valores superiores para preadolescentes (hombres $\alpha=.71$ y mujeres $\alpha=.67$) y adolescentes (hombres con $\alpha=.68$ y mujeres con $\alpha=.70$). Lo que se puede señalar finalmente es que aunque la consistencia interna de la escala *hostilidad* no ha sido la mejor en la adaptación colombiana, este valor es consistente con otras investigaciones, especialmente con versiones reducidas. Si se toma sólo el rango de mayor edad, la comparación mejora y las fiabilidades se acercan a las informadas en muestras españolas e italianas en versiones largas. Aquí también como en las anteriores escalas, se puede ver una mejoría de la consistencia interna en los grupos a medida que va incrementando la edad.

2.2.3 Consideraciones finales concernientes a la consistencia interna

Visto lo anterior, concluimos que esta versión reducida del AQ para preadolescentes y adolescentes de 20 ítems de Chahín-Pinzón *et al.* (2012) ostenta una adecuada consistencia interna para el total del instrumento, y para la escala *agresividad física*, en los tres grupos de edad estudiados en los que se ha dividido la muestra. Más no se puede afirmar lo mismo en las escalas *agresividad verbal* e *ira*, tanto para la muestra global como para los respectivos grupos de edad. No obstante lo afirmado, en el grupo de edad de 14 y 16 años la consistencia interna de las escalas es marginalmente aceptable y sus intervalos de confianza incluyen valores iguales o superiores a $\alpha=.70$, que pueden ser considerados adecuados. En consecuencia, estas

escalas pueden ser utilizadas en este rango de edad, teniendo el debido cuidado al realizar interpretaciones y sacar conclusiones.

Ahora, respecto al problema de la baja consistencia interna en la versión colombiana en las escalas *agresividad verbal*, *ira* y *hostilidad*, hay que tener en cuenta dos aspectos importantes, que se reflejan del análisis y comparaciones realizados. El primero, los valores obtenidos, aunque bajos, son consistentes con los estudios realizados hasta la fecha en población adolescente y preadolescentes. El segundo, la consistencia interna obtenida en el AQ en adultos tiende a decrecer sensiblemente cuando el instrumento se aplica en adolescentes, haciéndose mucho más crítica en preadolescentes Y tercero, la utilización de un menor número de ítems en el AQ, como se ha visto, generalmente disminuye la consistencia interna.

Con todo, queda la inquietud sobre cómo afrontar las consistencias de las escalas *agresividad verbal*, *ira* y *hostilidad* en edades menores a los 14 años. Aunque una posibilidad sería la aplicación de cuestionarios a padres y profesores (en vez de utilizar cuestionarios autoinformados), consideramos que otra alternativa válida sería llevar a cabo un estudio psicométrico para incrementar progresivamente el número de ítems del AQ entre los 8 y 14 años, hasta lograr una consistencia aceptable, tomando como referencia los indicadores de Santisteban *et al.* (2007). De esta forma se obtendría una nueva versión, con el menor número de ítems posible, que cumpliera con los requisitos mínimos respecto al coeficiente de consistencia interna. Obviamente con esto se perdería en parte el carácter de invariabilidad cultural logrado originalmente por Morales-Vives *et al.* (2005), pero hay que tener en cuenta que dicha

versión fue hecha sobre la base de muestras adultas, que tienen características psicométricas muy distintas a las de los adolescentes y preadolescentes.

Más no hay que perder de vista las desventajas metodológicas y prácticas que pueden tenerse aplicando versiones más largas en estas edades: el hecho que el tiempo se extienda puede llegar a ser un factor problemático a la hora de su aplicación en contextos escolares. Lo anterior es algo que debe ser evaluado cuidadosamente en el momento de seleccionar la longitud de la prueba, teniendo presente los pros y contras de cada alternativa en función de los objetivos y metodología para cada investigación.

2.3 Consideraciones sobre las diferencias de agresividad entre sexos

En lo concerniente a las diferencias debidas al sexo, los resultados conseguidos en la presente investigación informan un mayor valor de la escala *agresividad física* en varones ($F(1, 488)=29.15$; $p<.01$) con un tamaño del efecto de $d=.64$. El valor superior en esta escala ha sido reportado para adultos desde la creación del instrumento (Buss y Perry, 1992), y posteriormente en muestras con diversas características (Andreu, *et al.*, 2002; Fossati, *et al.*, 2003; Nakano, 2001; Ramírez, *et al.*, 2001; Von Collani y Werner, 2005). Asimismo, en niños y adolescentes el mayor valor de la escala *agresividad física* en varones ha sido informado por Morales-Vives (2007), Ang (2007), Reyna *et al.* (2011) y Santisteban *et al.* (2007). Se puede entonces afirmar que lo hallado en la muestra de niños colombianos es consistente con

un aspecto que ha sido reiterativo para el AQ en adultos y niños. Lo anterior lleva a aseverar que las diferencias en agresividad entre sexos de la muestra de Chahín-Pinzón *et al.* (2012) son más de tipo instrumental (como en este caso específico de la *agresividad física*) que de tipo cognitivo o emocional. Aspecto que según algunos estudios puede estar asociado a la manera en que culturalmente se espera que se exteriorice la conducta agresiva en los varones (Ang, 2007; Björkqvist, Lagerspetz y Kaukiainen, 1992; Lagerspetz, Björkqvist y Peltonen, 1988; Morales-Vives, 2007).

CONCLUSIONES

1 Conclusiones para la adaptación del test

BIS-11c

Las conclusiones de la adaptación de la Escala Barratt de Impulsividad para niños colombianos se enumeran a continuación:

- La adaptación colombiana del BIS-11c se ajusta al modelo de tres factores propuesto por Barratt (1985) y Cosi *et al.* (2008b) para medir la impulsividad.
- Hay equivalencia de los parámetros estructurales entre las soluciones factoriales de las versiones colombiana y española. Y la igualdad en el patrón de las cargas factoriales entre las dos versiones ha sido aceptada bajo ciertas restricciones.
- Las escalas *impulsividad motora* e *impulsividad no planificadora* de la adaptación colombiana presentan coeficientes de consistencia interna adecuados.
- La escala *impulsividad cognitiva* de la versión colombiana ha exhibido coeficientes de consistencia interna inaceptables.

2 Conclusiones para la adaptación del test AQ

Las conclusiones de la adaptación del Cuestionario de Agresividad de Buss y Perry para preadolescentes y adolescentes colombianos se enumeran a continuación:

- La adaptación colombiana replica la estructura tetrafactorial propuesta por Buss y Perry (1992) y Vigil-Colet *et. al.* (2005) para medir la agresividad.
- En la muestra la consistencia interna de la escala total del AQ posee coeficientes adecuados para todos los grupos de edad. La escala *agresividad física* y la escala *hostilidad* presentan coeficientes adecuados para todos los grupos de edad.
- En los grupos de 8 a 10 años y 11 a 13 años los coeficientes de consistencia interna de la escala *agresividad verbal* y la escala *ira* fueron inaceptables o estuvieron en el límite. Sólo en el grupo con mayor edad, de 14 a 16 años, se encontró que todas las escalas exhibieron coeficientes aceptables.
- Al comparar los puntajes de la escala *agresividad física* entre hombres y mujeres, los hombres han obtenido valores significativamente superiores.

REFERENCIAS

BIBIOGRÁFICAS

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV-TR)*. (4th ed.). Washington: American Psychiatric Association.

Andreu, J. M., Peña, M. E., y Graña, J. L. (2002). Adaptación psicométrica de la versión española del Cuestionario de Agresión. *Psicothema*, *14*(2), 476-482.

Ang, R. P. (2007). Factor structure of the 12-item aggression questionnaire: Further evidence from Asian adolescent samples. *Journal of Adolescence*, *30*(4), 671-685.

Arce, E., y Santisteban, C. (2006). Impulsivity: A review. *Psicothema*, *18*, 213-220.

Archer, J., Holloway, R., y McLoughlin, K. (1995a). Self-reported physical aggression among young men. *Aggressive Behavior*, *21*, 325-342.

Archer, J., Kilpatrick, G., y Bramwell, R. (1995b). A comparison of two aggression inventories. *Aggressive Behavior*, *21*, 371-380.

Asparouhov, T., y Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397-438.

- Balluerka, N., Gorostiaga, A., Alonso-Arbiol, I., y Haranburu, M. (2007). La adaptación de instrumentos de medida de unas culturas a otras: una perspectiva práctica. *Psicothema*, *19*, 124-133.
- Bandura, A. (1983). Psychological mechanisms of aggression. En Geen, R., y Donnerstein, E. (Eds.), *Aggression. Theoretical and Methodological Issues. Vol. I.* (pp. 1-40). New York: Academic Press.
- Barbero, M., Vila, E., y Holgado, F. P. (2008). La adaptación de los tests en estudios comparativos interculturales. *Acción Psicológica*, *5* (2), 7-16.
- Barratt, E. S. (1965). Factor analysis of some psychometric measures of impulsiveness and anxiety. *Psychological Reports*, *16*, 547-554.
- Barratt, E. S. (1985). Impulsiveness subtraits: Arousal and information processing. En Spence, J. T., e Izard, C. E. (Eds.), *Motivation, emotion and personality* (pp. 137-146). Amsterdam: Elsevier.
- Barratt, E. S. (1993). Impulsivity: Integrating cognitive, behavioral, biological and environmental data. En McCowan, W., y Shure, M. (Eds.), *The impulsive client: Theory, research and treatment* (pp. 39-56). Washington D.C.: American Psychological Association.

- Barratt, E. S. (1994). Impulsiveness and aggression. En Monahan J., y Steadman H. J. (Eds.), *Violence and mental disorder: developments in risk assessment* (pp. 61-79). Chicago: The University Chicago Press.
- Barratt, E. S., y Patton, J. H. (1983). Impulsivity: Cognitive, behavioral, and psychophysiological correlates. En Zuckerman M. (Ed.), *Biological bases of sensation seeking, impulsivity and anxiety* (pp. 77-121). Hillsdale, N. J.: Erlbaum.
- Barratt, E. S., y Slaughter, L. (1998). Defining, measuring, and predicting impulsive aggression: a heuristic model. *Behavioral Sciences & the Law*, 16(3), 285-302.
- Baylé, F. J., Bourdel, M. C., Caci, H., Gorwood, P., Chignon, J. M., Adés, J., y Lôo, H. (2000). Factor analysis of French translation of the Barratt impulsivity scale (BIS-10). *Canadian Journal of Psychiatry*, 45(2), 156-65.
- Bendig, A. W. (1962). Factor analytic scales of covert and overt hostility. *Journal of Consulting Psychology*, 26, 200-206.
- Bentler P. M., y Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.

Berkowitz, L. (1996). *Agresión. Causas, consecuencias y control*. Bilbao, España:

Editorial Desclée de Brouwer.

Bernstein, I. H., y Gesn, P. R. (1997). On the dimensionality of the Buss/Perry aggression questionnaire. *Behaviour Research and Therapy*, 35, 563-568.

Biaggio, M. K., Supplee, K., y Curtis, N. (1981). Reliability and validity of four anger scales. *Journal of Personality Assessment*, 45, 639-648.

Bishop, G. D., y Quah, S. H. (1998). Reliability and validity of measures of anger/hostility in Singapore: Cook & Medley Ho Scale, STAXI and Buss-Durkee Hostility Inventory. *Personality and Individual Differences*, 24, 867-878.

Björkqvist, K., Lagerspetz, K., y Kaukiainen, A. (1992). Do girls manipulate and boys fight? Developmental trends regarding direct and indirect aggression. *Aggressive Behavior*, 18, 117-127.

Bond, A., Verheyden, S., Wingrove, J., y Curran, H. V. (2004). Angry cognitive bias, trait aggression and impulsivity in substance users. *Psychopharmacology*, 171(3), 331-339.

- Browne, M. W. y Cudeck, R. (1993). Alternative ways of assessing model fit .En Bollen, K. A., y Long, J. S. (Eds.), *Testing Structural Equation Models* (pp. 136-162). Beverly Hills, C.A.: Sage.
- Brunas-Walgstaff, J., Tilley, A., Verity, M., Ford, S., y Thompson, D. (1997). Functional and dysfunctional impulsivity in children and their relationship to Eysenck's impulsiveness and venturesomeness dimensions. *Personality and Individual Differences*, 22, 19-25.
- Bryant, F. B., y Smith, B. D. (2001). Refining the architecture of aggression: A measurement model for the Buss-Perry Aggression Questionnaire. *Journal of Research in Personality*, 35, 138-167.
- Bushman, B. J., Cooper, H. M., y Lemke, H. M. (1991). Meta-analysis of factor analyses: An illustration using the Buss-Durkee Hostility Inventory. *Personality and Social Psychology Bulletin*, 17, 344-349.
- Buss, A. H., y Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology*, 21, 343-349.
- Buss, A. H., y Perry., M. P. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63, 452-459.

- Buss, A. H., y Plomin, R. (1975). *A temperament theory of personality development*.
New York: Wiley.
- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument:
A walk through the process. *Psicothema*, 20, 872-882.
- Carretero-Dios, H., y Pérez, C. (2005). Normas para el desarrollo y revisión de
estudios instrumentales. *International Journal of Clinical and Health
Psychology*, 5, 521-551.
- Castrillón, D. A., Ortiz, P. A., y Vieco, F. (2004). Cualidades paramétricas del
Cuestionario de Agresión (AQ) de Buss y Perry en estudiantes universitarios de
la ciudad de Medellín (Colombia). *Revista Facultad Nacional de Salud
Pública*, 22, 49-61.
- Chahín, N., Cosi, S., Lorenzo-Seva, U., y Vigil-Colet, A. (2010). Stability of the
factorial structure of Barrat's Impulsivity Scales for children across cultures: a
comparison between Spain and Colombia. *Psicothema*, 22, 983-989.
- Chahín, N. (2009). *Adaptación de la escala española de impulsividad para niños BIS-
11c a la población colombiana*. Trabajo fin de máster. Universitat Rovira i
Virgili.

- Chahín-Pinzón, N., Lorenzo-Seva, U., y Vigil-Colet, A. (2012). Características psicométricas de la adaptación colombiana del Cuestionario de Agresividad de Buss y Perry en una muestra de preadolescentes y adolescentes de Bucaramanga. *Universitas Psychologica, 11*(3), 979-988.
- Chung, Y. O., y Lee, C. W. (1997). A study of factor structures of the Barratt impulsiveness scale in Korean university students. *Korean Journal of Clinical Psychology, 16*, 117-129.
- Cloninger, C. R., Przybeck, T. R., y Svrakic, D. M. (1991). The tridimensional personality questionnaire: U.S. normative data. *Psychological Reports, 69*, 1047-1057.
- Condon, L., Morales-Vives, F., Ferrando, P. J., y Vigil-Colet, A. (2006). Sex Differences in the full and reduced versions of the Aggression Questionnaire: A question of differential item. *European Journal of Psychological Assessment, 22*(2), 92-97.
- Cosi, S. (2008). *Evaluación de la impulsividad infantil: Adaptación del Barratt Impulsivity Scale*. Trabajo fin de máster. Universitat Rovira i Virgili.
- Cosi, S. (2011). Evaluación de la impulsividad y su papel como factor predictor de psicopatología emocional en población escolar. Tesis doctoral. Universitat Rovira i Virgili.

- Cosi, S., Hernández-Martínez, C., Canals, J., y Vigil-Colet A. (2011). Impulsivity and internalizing disorders in childhood. *Psychiatry Research*, 190, 342-347.
- Cosi, S., Morales-Vives, F., Canals, J., Lorenzo-Seva, U., y Vigil-Colet, A. (2008a). Functional and dysfunctional impulsivity in childhood and adolescence. *Psychological Reports*, 103, 67-76.
- Cosi, S., Vigil-Colet, A., Canals, J. (2008b). *Desarrollo de la versión mejorada del BIS-11c: Relaciones con agresividad y rendimiento académico*. Póster presentado en X Jornadas de la Sociedad Española para la Investigación de las Diferencias Individuales, Salamanca, España.
- Cosi, S., Vigil-Colet, A., Canals, J., y Lorenzo-Seva, U. (2008c). Psychometric properties of the Spanish Adaptation of the Barratt Impulsiveness Scale-11a for children. *Psychological Reports*, 103, 336-346.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(2), 297-334.
- Diamond, P. M., Wang, E. W., y Buffington-Vollum, J. (2005). Factor structure of the Buss-Perry Aggression Questionnaire (BPAQ) with mentally ill male prisoners. *Criminal Justice and Behavior*, 32(5), 546-564.

- Dickman, S. J. (1990). Functional and dysfunctional impulsivity, personality and cognitive correlates. *Journal of Personality and Social Personality*, 58, 95-102.
- Dom, G., D'haene, P., Hulstijn, W., y Sabbe, B. (2006). Impulsivity in abstinent early-and late-onset alcoholics: differences in self-report measures and a discounting task. *Addiction*, 101(1), 50-59.
- Dougherty, D. M., Mathias, C. W., Marsh, D. M., Papageorgiou, T. D., Swann, A. C., y Moeller, F. G. (2004). Laboratory measured behavioral impulsivity relates to suicide attempt history. *Suicide and Life Threatening Behavior*, 34, 374-385.
- Edmunds, G., y Kendrick, D. C. (1980). *The measurement of human aggressiveness*. New York: Wiley.
- Ercikan, K. (2002). Disentangling sources of differential item functioning in multilanguage assessments. *International Journal of Testing*, 2(3), 199-215.
- Esteban, A., y Taberner, C. (2011). Relationship between impulsiveness and deviant behavior among adolescents in the classroom: Age and sex differences. *Psychological Reports*, 109, 703-717.
- Evenden, J. L. (1999). Varieties of impulsivity. *Psychopharmacology*, 146(4), 348-361.

Eysenck, H. J., y Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum.

Eysenck, S. B., y Eysenck, H. J. (1977). The place of impulsiveness in a dimensional system of personality description. *British Journal of Social and Clinical Psychology, 16*, 57-68.

Eysenck, S. B., Pearson, P. R., Easting, G., y Allsopp, J. F. (1985). Age norms for impulsiveness, venturesomeness and empathy in adults. *Personality and Individual Differences, 6*(5), 613-619.

Eysenck, S. B., y Eysenck, H. J. (1978). Impulsiveness and venturesomeness: Their position in a dimensional system of personality description. *Psychological reports, 43*, 1247-1255.

Fan, X., y Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research, 42*, 509-529.

Felsten, G., y Leitten, C. L. (1993). Expressive, but not neurotic hostility is related to cardiovascular reactivity during a hostile competitive task. *Personality and Individual Differences, 14*, 805-813.

Fink, A. D., y McCown, W. G. (1993). Impulsivity in children and adolescents: Measurement, causes and treatment. En McCown, W. G., Johnson, J. L., y

- Shure, M. B. (Eds.), *The impulsive client, theory, research and treatment* (pp. 279-308). Washington: American Psychological Association.
- Folino, J. O., Escobar-Córdoba, F., y Castillo, J. L. (2006). Exploración de la validez de la Escala de Impulsividad de Barratt en la población carcelaria argentina. *Revista Colombiana de Psiquiatría*, 35, 132-148.
- Forcada, R., Pardo, N., y Bondia, B. (2006). Impulsividad en dependientes de cocaína que abandonan el consumo. *Adicciones*, 18, 111-118.
- Fossati, A., Barratt, E. S., Acquarini, E., y Di Ceglie, A. (2002). Psychometric properties of an adolescent version of the Barratt Impulsiveness Scale-11 for a sample of Italian high school students. *Perceptual and Motor Skills*, 95, 621-635.
- Fossati, A., Di Ceglie, A., Acquarini, E., y Barratt, E. S. (2001). Psychometric properties of an Italian version of the Barratt Impulsiveness Scale-11 (BIS-11) in nonclinical subjects. *Journal of Clinical Psychology*, 57, 815-828.
- Fossati, A., Maffei, C., Acquarini, E., y Di Ceglie, A. (2003). Multigroup confirmatory component and factor analyses of the Italian version of the Aggression Questionnaire. *European Journal of Psychological Assessment*, 19(1), 54-65.

- Gallardo-Pujol, D., Krarnp, U., Garcia-Forero, C., Perez-Ramirez, M., y Andres-Pueyo, A. (2006). Assessing aggressiveness quickly and efficiently: the Spanish adaptation of Aggression Questionnaire-Refined version. *European Psychiatry, 21*(7), 487-494.
- Garcia-León, A., Reyes, G. A., Vila, J., Pérez, N., Robles, H., y Ramos, M. M. (2002). The aggression questionnaire: A Validation study in student samples. *The Spanish Journal of Psychology, 5*, 45-53.
- Gerbing, D. W., Ahadi, S. A. y Patton, J. H. (1987). Toward a conceptualization of impulsivity: Components across the behavioral and self-report domains. *Multivariate Behavioral Research, 22*, 357-379.
- Gilbert, K. E., Kalmar, J. H., Womer, F. Y., Markovich, P. J., Pittman, B., Nolen-Hoeksema, S., y Blumberg, H. P. (2011). *Impulsivity in adolescent bipolar disorder. Acta Neuropsychiatrica, 23*, 57-61.
- Giotakos, O., Markianos, M., Vaidakis, N., y Christodoulou, G. N. (2003). Aggression, impulsivity, plasma sex hormones, and biogenic amine turnover in a forensic population of rapists. *Journal of Sex and Marital Therapy, 29*, 215-225.
- Goudriaan, A. E., Oosterlaan, J., De Beurs, E., y Van den Brink, W. (2008). The role of self-reported impulsivity and reward sensitivity versus neurocognitive

measures of disinhibition and decision making in the prediction of relapse in pathological gamblers. *Psychological Medicine*, 38, 41-50.

Güleç, H., Tamam, L. Güleç, M. Y., Turhan, M., Karakusl, G., Zengin, M., y Stanford, M .S. (2008). Psychometric properties of the Turkish version of the Barratt Impulsiveness Scale. *Bulletin of Clinical Psychopharmacology*, 18, 251-258.

Gunn, J., y Gristwood, J. (1975). Use of the Buss-Durkee Hostility Inventory among British prisoners. *Journal of Consulting and Clinical Psychology*, 43, 590.

Haden, S. C., y Shiva, A. (2008). Trait impulsivity in a forensic inpatient sample: An evaluation of the Barratt Impulsiveness Scale. *Behavioral Sciences and the Law*, 27, 675-690.

Hambleton, R. K. (2001). The next generation of ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment*, 17, 164-172.

Hambleton, R. K. (2005). Issues, designs and technical guidelines for adapting tests into multiple languages and cultures. En Hambleton, R. K., Merenda, P. F., y Spielberger, C. (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). London: L.E.A

- Hambleton, R. K., y Patsula, L. (1998). Adapting tests for use in multiple languages and cultures. *Social Indicators Research*, 43, 153-171.
- Hambleton, R. K., y Zenisky, A. L. (2011). Translating and adapting tests for cross cultural assessments. En Matsumoto, D. y Van de Vijver, F. J. R. (Eds.), *Cross-Cultural Research Methods in Psychology* (pp. 46-70). New York: Cambridge University Press.
- Harris, J. A. (1995). Confirmatory factor analysis of the Aggression Questionnaire. *Behaviour Research and Therapy*, 33(8), 991-993.
- Harris, J. A. (1997). A further evaluation of the aggression questionnaire: issues of validity and reliability. *Behaviour Research and Therapy*, 35, 1047-1053.
- Hartmann, A. S., Rief, W., y Hilbert, A. (2011). Psychometric properties of the German version of the Barratt Impulsiveness Scale, Version 11 (BIS-11) for adolescents. *Perceptual and Motor Skills*, 112, 353-368.
- Henry, C., Mitropoulou, V., New, A. S., Koenigsberg, H. W., Silverman, J., y Siever, L. J. (2001). Affective instability and impulsivity in borderline personality and bipolar II disorders: similarities and differences. *Journal of Psychiatric Research*, 35, 307-312.

- Holland, P. W., y Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. En Wainer, H., y H. I. Braun, H. I. (Eds.), *Test validity* (pp. 129-145). Hillsdale, N.J.: Lawrence Erlbaum.
- Hornsveld, R. H. J., Muris, P., Kraaimaat, F. W., y Meesters, C. (2009). The Aggression Questionnaire in Dutch violent forensic psychiatric patients and secondary vocational students. *Assessment, 16*, 181-192.
- Hou, L., Xiao, S., He, X., Li, J., y Lui, H. (2006). Reliability and validity of Chinese version of Barratt Impulsiveness Scale-11. *Chinese Journal of Clinical Psychology, 4*, 343-345.
- Hu, L. T., y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6* (1), 1-55.
- Hu, L., y Bentler, P. M. (1995). Evaluating model fit. En Hoyle. R. (Ed.), *Structural equation modelling: Concepts, issues and applications* (pp.76-99). Thousand Oaks, C.A.: Sage Publications.
- Hui, C. H., y Triandis, H. C. (1985). Measurement in cross cultural psychology. *Journal of Cross-Cultural Psychology, 16*, 131-152.

- Ireland, J. L. y Archer, J. (2008). Impulsivity among adult prisoners: A confirmatory factor analysis study of the Barratt Impulsivity Scale. *Personality and Individual Differences, 45*(4), 286-292.
- Lagerspetz, K., Björkqvist, K., y Peltonen, T. (1988). Is indirect aggression typical of females? Gender differences in aggressiveness in 11 to 12 years-old children. *Aggressive Behavior, 14*, 403-414.
- Lane, S. D., Moeller, F. G., Steinberg, J. L., Buzby, M., y Kosten, T. R. (2007). Performance of cocaine dependent individuals and controls on a response inhibition task with varying levels of difficulty. *American Journal of Drug and Alcohol Abuse, 33*, 717-726.
- Lejuez, C. W., Bornovalova, M. A., Reynolds, E. K., Daughters, S. B., y Curtin, J. J. (2007). Risk factors in the relationship between gender and crack/cocaine. *Experimental and Clinical Psychopharmacology, 15*, 165-175.
- Leshem, R., y Glicksohn, J. (2007). The construct of impulsivity revisited. *Personality and Individual Differences, 43*, 681-692.
- Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*, 53-76.

- López del Pino, M. C., Sánchez, A., Rodríguez, L. T., y Fernández M. P. (2009).
Propiedades psicométricas del cuestionario AQ aplicado a población
adolescente. *EduPsykhé*, 8, 79-94.
- Lorenzo-Seva, U. (2003). A factor simplicity index. *Psychometrika*, 68, 49-60.
- Lorenzo-Seva, U., Kiers, H. A. L., y Ten Berge, J. M. F. (2002). Techniques for
oblique factor rotation of two or more loading matrices to a mixture of simple
structure and optimal agreement. *British Journal of Mathematical and
Statistical Psychology*, 55, 337-360.
- Lorenzo-Seva, U., y Ferrando, P. J. (2006). FACTOR: A computer program to fit the
exploratory factor analysis model. *Behavioral Research Methods, Instruments
and Computers*, 38, 88-91.
- Loyola, N. (2011). *Validez y confiabilidad de la escala de impulsividad de Barratt
versión 11 (BIS-11) en mujeres encarceladas*. Tesis de licenciatura.
Universidad Católica del Perú.
- Luengo, M. A., Carrillo de la Peña, M. T. y Otero, J. M. (1991). The components of
impulsiveness: a comparison of the I.7 Impulsiveness Questionnaire and the
Barratt impulsiveness Scale. *Personality and Individual Differences*, 12, 657-
667.

- Malloy-Diniz, L., Fuentes, D., Leite, W. B., Correa, H., y Bechara, A. (2007). Impulsive behavior in adults with attention deficit/hyperactivity disorder: Characterization of attentional, motor and cognitive impulsiveness. *Journal of the International Neuropsychological Society*, *13*, 693-698.
- Matthews, K. A., y Saal, F. E. (1978). Relationship of the Type A coronary-prone behavior pattern to achievement, power, and affiliation motives. *Psychosomatic Medicine*, *40*, 630-636.
- Maxwell, J. P. (2007). Development and preliminary validation of a Chinese version of the Buss-Perry Aggression Questionnaire in a population of Hong Kong Chinese. *Journal of personality assessment*, *88*(3), 284-294.
- Meesters, C. M., Muris, P., Bosma, H., Schouten, E., y Beuving, S. (1996). Psychometric evaluation of the Dutch version of the Aggression Questionnaire. *Behaviour Research and Therapy*, *34*, 839-843.
- Meule, A., Vögele, C., y Kübler, A. (2011). Psychometric evaluation of the German Barratt Impulsiveness Scale - Short version (BIS-15). *Diagnostica*, *57*, 126-13.
- Millsap, R. E., y Meredith, W. (2007). Factorial invariance: Historical perspectives and new problems. En Cudeck, R., y MacCallum, R. C. (Eds.), *Factor analysis at 100: Historical developments and new directions* (pp 131-152). Mahwah, N.J.: Erlbaum.

Moeller, F. G., Dougherty, D. M., Barratt, E. S., Schmitz, J. M., Swann, A. C., y Grabowski, J. (2001). The impact of impulsivity on cocaine use and retention in treatment. *Journal of Substance Abuse and Treatment*, 21, 193-198.

Moeller, F. G., Barratt, E. S., Dougherty, D. M., Schmitz, J. M., y Swann, A. C. (2001). Psychiatric aspects of impulsivity. *American Journal of Psychiatry*, 158, 1783-1793.

Mook, J., Kleijn, W. C., y Van der Ploeg, H. M. (1991). Symptom-positively and negatively worded items in two popular self-report inventories of anxiety and depression. *Psychological Reports*, 69, 551-560.

Morales-Vives, F (2007). *El efecto de la impulsividad sobre la agresividad y sus consecuencias en el rendimiento de los adolescentes*. Tesis doctoral. Universitat Rovira i Virgili.

Morales-Vives, F., Codorniu-Raga, M. J., y Vigil-Colet, A. (2005). Características psicométricas de las versiones reducidas del cuestionario de agresividad de Buss y Perry. *Psicothema*, 17(1), 96-100.

Morgan, J. E., Gray, N. S., y Snowden, R. J. (2011). The relationship between psychopathy and impulsivity: A multi-impulsivity measurement approach. *Personality and Individual Differences*, 51(4), 429-434.

- Morren, M., y Meesters, C. (2002). Validation of the Dutch version of the Aggression Questionnaire in adolescent male offenders. *Aggressive Behavior*, 28, 87-96.
- Muñiz, J. (1994). Traducción/Adaptación de tests educativos y psicológicos. *Papeles del Psicólogo*, 59, 43-44.
- Muñiz, J., y Hambleton, R. K. (1996). Directrices para la traducción y adaptación de los tests. *Papeles del Psicólogo*, 66, 63-70.
- Muthen, L. K., y Muthen, B. O. (2008). Mplus (Version 5.1). Los Angeles, C.A.: Muthen & Muthen.
- Nakano, K. (2001). Psychometric evaluation on the Japanese adaptation of the aggression questionnaire. *Behavioral Research and Therapy*, 39, 853-858.
- Nandagopal, J. J., Fleck, D. E., Adler, C. M., Mills, N. P., Strakowski, S. M., y Del Bello, M. P. (2011). Impulsivity in adolescents with bipolar disorder and/or attention deficit/hyperactivity disorder and healthy controls as measured by the Barratt Impulsiveness Scale. *Journal of Child and Adolescent Psychopharmacology*, 21, 465-468.
- Oquendo, M. A., Baca-Garcia, E., Graver, R., Morales, M., Montalvan, V., y Mann, J. J. (2001). Spanish adaptation of the Barratt impulsiveness scale (BIS-11). *European Journal of Psychiatry*, 15(3), 147-55.

- Orozco-Cabal, L., Rodríguez, M., Herin, D.V., Gempeler J., y Uribe, M. (2010).
Validity and reliability of the abbreviated Barratt Impulsiveness Scale in
Spanish (BIS-15S). *Revista Colombiana de Psiquiatría*, 39(1), 93-109.
- Parker, J. D. A., y Bagby, R .M. (1997). Impulsivity in adults: a critical review of
measurement approaches. En Webster, C. D., y Jackson, M. A. (Eds.),
Impulsivity: theory, assessment, and treatment (pp. 142–157). New York: The
Guilford Press.
- Patton, J. H., Stanford, M. S., y Barratt, E. S. (1995). Factor structure of the Barratt
Impulsiveness Scale. *Journal of Clinical Psychology*, 51, 768-774.
- Patton, J. H., y Stanford, M. S. (2011). Psychology of impulsivity. En Grant, J., y
Potenza, M. (Eds.), *The Oxford handbook of impulse control disorders* (pp.
262–275). New York: Oxford.
- Peluso, M. A. M., Hatch, J. P., Glahn, D. C., Monkul, E. S., Sanches, M., Najt, P.,
Bowden, C. L., Barratt, E. S., y Soares, J. C. (2007). Trait impulsivity in
patients with mood disorders. *Journal of Affective Disorders*, 100, 227-231.
- Petzel, T. P., y Michaels, E. J. (1973). Perception of violence as a function of levels of
hostility. *Journal of Consulting and Clinical Psychology*, 41, 35-36.

- Porras, S., Salamero, M., y Sender, R. (2001). Adapatación española del Buss-Perry Aggression Questionnaire. *Cuadernos de Medicina Psicosomática y Psiquetría de Enlace*, 60, 7-12.
- Preuss, U. W., Rujescu, D., Giegling, I., Watzke, S., Koller, G., Zetsche, T., Meisenzalhal, E. M., Soyca, M., y Möller, H. J. (2008). Psychometric evaluation of the German version of the Barratt Impulsiveness Scale. *Nervenarzt*, 79, 305-319.
- Prochazka, H., y Agren, H. (2001). Aggression in the general Swedish population, measured with a new self-rating inventory: The Aggression Questionnaire-Revised Swedish version (AQ-RSW). *Nordic Journal of Psychiatry*, 55, 17-23.
- Ramírez, J., Andreu, J. M., y Fujihara, T. (2001). Cultural and sex differences in aggression: A comparison between Japanese and Spanish students using two different inventories. *Aggressive Behavior*, 27, 313-322.
- Rapee, R. M., Barrett, P. M., Dadds, M. R., y Evans, L. (1994). Reliability of the DSM-III-R childhood anxiety disorders using structured interview: Interrater and parent-child agreement. *Journal of the American Academy of Child and Adolescent Psychiatry*, 33, 984-992.
- Real Academia Española. (2001). *Diccionario de la lengua española* (22.^a edición). Madrid: Espasa Calpe.

- Recio, P. (2012). *Equivalencia e invarianza de medida entre grupos: análisis factorial confirmatorio vs teoría de respuesta al ítem*. Tesis doctoral. Universidad Complutense de Madrid.
- Reyna, C., Lello, M.G., Sanchez, A., y Brussino, S. (2011). The Buss-Perry Aggression Questionnaire: Construct validity and gender invariance among Argentinean adolescents. *International Journal of Psychological Research*, 4(2), 30-37.
- Rodriguez-Jimenez, R., Avila, C., Jimenez-Arriero, M. J., Ponce, G., Monasor, R., Jiménez, M., Aragües, M., Hoenicka, J., Rubio, G., y Palomo, T. (2006). Impulsivity and sustained attention in pathological gamblers: influence of childhood ADHD history. *Journal of Gambling Studies*, 22, 451-461.
- Santisteban, C., Alvarado, J. M., y Recio, P. (2007). Evaluation of a Spanish version of the Buss and Perry Aggression Questionnaire: Some personal and situational factors related to the aggression scores of young subjects. *Personality and Individual Differences*, 42, 1453-1462.
- Santisteban, C., y Alvarado, J. M. (2009). The Aggression Questionnaire for Spanish preadolescents and adolescents: PA-AR. *The Spanish Journal of Psychology*, 12, 320-326.

- Shepard, L., Camelli, G., y Averill, M. (1981). Comparison of six procedures for detecting test item bias using both internal and external ability criteria. *Journal of Educational Statistics*, 6, 317-375.
- Siegmán, A. W., Dembroski, T. M., y Ringel, N. (1987). Components of hostility and the severity of coronary artery disease. *Psychosomatic Medicine*, 49, 127-135.
- Sierra, J. C., y Gutiérrez, R. (2007). Validación de la versión española del Cuestionario de Agresión de Buss-Perry en estudiantes universitarios salvadoreños. *Psicología y Salud*, 17, 103-113.
- Singh, P., Solanki, R. K., y Bhatnagar, P. S. (2008). BIS-11A Hindi version: A preliminary study of impulsivity in rural and urban Indian adolescents. *Indian Journal of Psychiatry*, 50, 96-99.
- Sireci, S. G. (2011). Evaluating test and survey items for bias across languages and cultures. En Matsumoto, D., y Van de Vijver, F. J. R. (Eds.), *Cross-cultural research methods in psychology* (pp. 216-240). New York: Cambridge University Press.
- Skinner, M. D., Aubin, H. J., y Berlin, I. (2004). Impulsivity in smoking, nonsmoking, and ex-smoking alcoholics. *Addictive Behaviors*, 29(5), 973-978.

- Smith, P., Waterman, M., y Ward, N. (2006). Driving aggression in forensic and non-forensic populations: Relationships to self-reported levels of aggression, anger and impulsivity. *British Journal of Psychology*, 97(3), 387-403.
- Someya, T., Sakado, K., Seki, T., Kojima, M., Reist, C., Tang, S.W., y Takahashi, S. (2001). The Japanese version of the Barratt Impulsiveness Scale, 11th version (BIS-11): its reliability and validity. *Psychiatry and Clinical Neurosciences*, 55(2), 111-114.
- Sommantico, M., Osorio, M., Parrello, S., De Rosa, B., y Donizzetti, A. R. (2008). Validación de la versión italiana del Cuestionario de Agresión (AQ) en el sur de Italia. *Revista Electrónica de Psicología Iztacala*, 1, 28-45.
- Spinella, M. (2007). Normative data and a short form of the Barratt Impulsiveness Scale. *International Journal of Neuroscience*, 117, 359-368.
- Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L., Anderson, N. E., y Patton, J. M. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences*, 47, 385-395.
- Suris, A., Lind, L., Emmetta, G., Bormanc, P. D., Kashner, M., y Barratt, E. S. (2004). Measures of aggressive behavior: Overview of clinical and research instruments. *Aggression and Violent Behavior*, 9, 165-227.

Swann, A. C., Pazzaglia, P., Nicholls, A., Dougherty, D. M., y Moeller, F. G. (2003).

Impulsivity and phase of illness in bipolar disorder. *Journal of Affective Disorders*, 73, 105-111.

Swann, A. C., Steinberg, J. L., Lijffijt, M., y Moeller, F. G. (2008). Impulsivity:

Differential relationship to depression and mania in bipolar disorder. *Journal of Affective Disorders*, 106, 241-248.

Tsorbatzoudis, H. (2006). Psychometric evaluation of the Greek version of the

Aggression Questionnaire. *Perceptual and Motor Skills*, 102, 703-718.

Tucker, L. R. (1951). A method for synthesis of factor analysis studies. *Personnel*

Research Section Report N° 984. Washington D.C.: Department of the Army.

Van de Vijver, F. J. R., y Leung, K. (2000). Methodological issues in psychological

research on culture. *Journal of Cross-Cultural Psychology*, 31, 33-51.

Van de Vijver, F. J. R., y Poortinga, Y. H. (1997). Towards an integrated analysis of

bias in cross-cultural assessment. *European Journal of Psychological Assessment*, 13, 29-37.

Van de Vijver, F. J. R., y Poortinga, Y. H. (2005). Conceptual and methodological

issues in adapting tests. En Hambleton, R. K., Merenda, P. F., y Spielberger, C.

- (Eds.), *Adapting educational and psychological tests for cross-cultural Assessment* (pp. 39-64). London: L.E.A.
- Van de Vijver, F., y Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: an overview. *European Review of Applied Psychology*, 54, 119-135.
- Van de Vijver, F., y Lonner, W. (1995). A bibliometric analysis of the Journal of Cross-Cultural Psychology. *Journal of Cross-Cultural Psychology*, 26, 591-602.
- Van de Vijver, F. J. R., y Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1, 89-99.
- Vasconcelos, A. G., Malloy-Diniz, L., y Correa, H. (2012). Systematic review of psychometric properties of Barratt impulsiveness scales version 11 (BIS-11). *Clinical Neuropsychiatry*, 9, 61-74.
- Vigil-Colet, A., Lorenzo-Seva, U., Codorniu-Raga, M. J., y Morales, F. (2005). Factor structure of the aggression questionnaire among different samples and languages. *Aggressive Behavior*, 31, 601-608.

- Vitoratou, S., Ntzoufras, I., Smyrnis, N., y Stefanis, N. C. (2009). Factorial composition of the Aggression Questionnaire: A multi-sample study in Greek adults. *Psychiatry Research*, *168*(1), 32-39.
- Von Collani, G., y Werner, R. (2005). Self-related and motivational constructs as determinants of aggression: An analysis and validation of a German version of the Buss–Perry Aggression Questionnaire. *Personality and Individual Differences*, *38*(7), 1631-1643.
- Von Diemen L., Szobot, C. M., y Kessler, F., y Pechansky, F. (2007). Adaptation and construct validation of the Barratt Impulsiveness Scale (BIS 11) to Brazilian Portuguese for use in adolescents. *Revista Brasileira de Psiquiatria*, *29*(2), 153-156.
- Whiteside, S. P., y Lynam, D. R. (2001). The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences*, *30*, 669-689.
- Williams, T. Y., Boyd, J. C., Cascard, I. M. A., y Poythress, N. (1996). Factor structure and convergent validity of the Aggression Questionnaire in an offender population. *Psychological Assessment*, *8*(4), 398-403.

Yang, H .Q., Yao, S. Q., y Zhu, X.. Z. (2007). The Chinese version of the Barratt impulsiveness scale 11th version (BIS-11) in college students: Its reliability and validity. *Chinese Mental Health Journal*, 21, 223-225.

Yao, S., Yang, H., Zhu, X., Auerbach, R. P., Abela, J. R., Pulleyblank, R. W., y Tong, X. (2007). An examination of the psychometric properties of the Chinese version of the Barratt Impulsiveness Scale, 11th version in a sample of Chinese adolescents. *Perceptual and Motor Skills*, 104, 1169-82.

Zaparniuk, J., y Taylor, S. (1997). Impulsivity in children and adolescents. En Webster, C. D., y Jackson, M. A. (Eds.), *Impulsivity: Theory, assessment, and treatment* (pp.158-179). New York: Guilford.