

Cuestionario de Agresión (AQ) en estudiantes mexicanos de bachillerato: evidencias de validez estructural y confiabilidad

Aggression Questionnaire (AQ) in Mexican high school students: evidence of structural validity and reliability

Questionário de Agressão (AQ) em estudantes mexicanos do ensino médio: evidências de validade estrutural e confiabilidade

Alejandro César Antonio Luna Bernal

Universidad de Guadalajara, Centro Universitario de Ciencias Sociales y Humanidades,

Departamento de Filosofía, México

alejandro.luna@academicos.udg.mx

<https://orcid.org/0000-0002-3371-5999>

Resumen

El estudio de la agresión resulta relevante para comprender problemáticas relacionadas con la violencia y el conflicto en adolescentes. A nivel internacional, el Cuestionario de Agresión (AQ) es uno de los instrumentos más utilizados para estudiar esta variable; sin embargo, no se identificaron estudios sobre sus propiedades psicométricas en adolescentes mexicanos. Por ello, el presente estudio se propuso evaluar la validez estructural y la confiabilidad de este instrumento en una muestra de adolescentes mexicanos de un plantel educativo de nivel medio superior. Participaron 683 estudiantes de un plantel público ubicado en el área metropolitana de Guadalajara, Jalisco, México, quienes respondieron la versión de 29 ítems del instrumento. Se realizó un análisis factorial confirmatorio y se estimó la consistencia interna. Los resultados respaldaron como parcialmente aceptable un modelo revisado de 26 ítems con cuatro factores subyacentes correlacionados: agresión física, agresión verbal, ira y hostilidad, en concordancia con el modelo teórico original. Se obtuvieron índices de bondad de ajuste parcialmente aceptables ($\chi^2/gl = 4.64$, CFI = .933, TLI = .925, RMSEA = .073 (IC 90% [.069, .077]), SRMR = .056). Los coeficientes de consistencia interna resultaron adecuados en las escalas de agresión verbal, hostilidad e ira ($>.70$), buenos en agresión física

(> .80), y excelentes para el instrumento total (> .90). En conclusión, el estudio presenta adecuadas estimaciones de confiabilidad del instrumento, pero evidencias de validez estructural parcialmente aceptables, por lo que la interpretación de sus puntuaciones debe realizarse con cautela al estudiar la agresión disposicional en adolescentes desde una perspectiva multidimensional.

Palabras clave: adolescencias, agresión, hostilidad, ira, psicometría.

Abstract

The study of aggression is important to understanding problems related to violence and conflict in adolescents. The Aggression Questionnaire (AQ) is one of the most widely used instruments internationally to study this variable; however, no studies on its psychometric properties in Mexican adolescents were identified. Therefore, the aim of this study was to evaluate the structural validity and reliability of the AQ in a sample of Mexican adolescents from a high school educational institution. A total of 683 students from a public high school located in Guadalajara metropolitan area, Jalisco, Mexico, participated and completed the 29-item version of the AQ. Confirmatory factor analysis and internal consistency analyses were conducted. The results partially supported as acceptable a revised 26-item model with four correlated underlying factors: physical aggression, verbal aggression, anger, and hostility, in line with the original theoretical model. Partially acceptable goodness-of-fit indices were obtained ($\chi^2/df = 4.64$, CFI = .933, TLI = .925, RMSEA = .073 (IC 90% [.069, .077]), SRMR = .056). Internal consistency was adequate for the verbal aggression, hostility and anger scales (> .70), satisfactory for physical aggression (> .80), and excellent for the overall instrument (> .90). In conclusion, the study provides adequate reliability estimates for the instrument but only partially acceptable evidence of structural validity; therefore, its scores should be interpreted with caution when studying dispositional aggression in adolescents from a multidimensional perspective.

Keywords: adolescence, aggression, hostility, anger, psychometrics.

Resumo

O estudo da agressão é relevante para compreender questões relacionadas à violência e ao conflito entre adolescentes. Internacionalmente, o Questionário de Agressão (AQ) é um dos instrumentos mais utilizados para estudar essa variável; no entanto, não foram identificados estudos sobre suas propriedades psicométricas em adolescentes mexicanos. Por isso, o presente estudo propôs-se a avaliar a validade estrutural e a confiabilidade desse instrumento em uma amostra de adolescentes mexicanos de uma instituição de ensino de nível médio superior. Participaram 683 estudantes de uma escola pública localizada na região metropolitana de Guadalajara, Jalisco, México, que responderam à versão de 29 itens do instrumento. Realizou-se uma análise fatorial confirmatória e estimou-se a consistência interna. Os resultados corroboraram, como parcialmente aceitável, um modelo revisado de 26 itens com quatro fatores subjacentes correlacionados: agressão física, agressão verbal, raiva e hostilidade, em consonância com o modelo teórico original. Foram obtidos índices de qualidade de ajuste parcialmente aceitáveis ($\chi^2/gl = 4,64$; CFI = 0,933; TLI = 0,925; RMSEA = 0,073 [IC 90%: 0,069; 0,077]; SRMR = 0,056). Os coeficientes de consistência interna mostraram-se adequados nas escalas de agressão verbal, hostilidade e raiva ($> 0,70$), bons na de agressão física ($> 0,80$) e excelentes para o instrumento como um todo ($> 0,90$). Em conclusão, o estudo apresenta estimativas adequadas de confiabilidade do instrumento, mas evidências de validade estrutural parcialmente aceitáveis; portanto, a interpretação de suas pontuações deve ser feita com cautela ao estudar a agressão disposicional em adolescentes sob uma perspectiva multidimensional.

Palavras-chave: adolescência, agressão, hostilidade, raiva, psicometria.

Fecha Recepción: Enero 2026

Fecha Aceptación: Julio 2026

Introducción

El estudio de la agresión en adolescentes resulta relevante para comprender procesos mediante los cuales se producen y desarrollan dinámicas de violencia y conflicto en esta etapa vital (Rodríguez Molinero & Imaz Roncero, 2020); por ello, el conocimiento derivado de ese estudio es fundamental para el diseño de estrategias de formación, prevención y atención en el marco de la educación para la paz, orientada a promover una cultura de paz fundada en el respeto de los derechos humanos y en los Objetivos de Desarrollo Sostenible (Jiménez Velázquez, 2025).

Además, un mayor conocimiento de la agresión es relevante debido a su estrecha relación con múltiples aspectos clave del desarrollo adolescente (Pomahuacre-Carhuayal et al., 2024). De acuerdo con Javela et al. (2023), la adolescencia es una etapa marcada por transformaciones cerebrales que inciden en la regulación emocional y la toma de decisiones, etapa en que las conductas agresivas se han asociado a problemas clínicos, al desarrollo de trayectorias antisociales y a consecuencias como el abandono escolar y el desempleo. Por ello, la identificación precisa de la agresión podría ser relevante para el diseño de intervenciones eficaces.

En ese marco, resulta fundamental contar con instrumentos psicométricos con evidencias de validez y estimaciones adecuadas de confiabilidad para el estudio de la agresión.

El Cuestionario de Agresión (*Aggression Questionnaire*, AQ) de Buss y Perry (1992) es probablemente el instrumento más conocido a nivel mundial para medir la agresión disposicional mediante autoinforme (Krahé, 2021). De acuerdo con Gerevich et al. (2007), una vez que se publicó en 1992, el AQ se convirtió rápidamente en un estándar de referencia para la medición de la agresión. Una de las ventajas que tiene este instrumento es el estar constituido por varias escalas (hostilidad, ira, agresión física y agresión verbal), lo que permite obtener una estimación multidimensional del grado de agresión disposicional que presentan los individuos. De acuerdo con Chester et al. (2024), esta perspectiva que concibe la agresión como integrada por componentes de tipo conductual, afectivo y cognitivo constituye actualmente el enfoque más ampliamente adoptado para el estudio de la agresión considerada como rasgo o característica de personalidad.

En este modelo propuesto por Buss y Perry (1992), la agresión entendida como rasgo de personalidad se considera un constructo multidimensional integrado por cuatro componentes. La agresión física y verbal constituyen el componente conductual, la ira el componente afectivo, y la hostilidad el componente cognitivo.

La agresión física se refiere a una propensión a golpear a otras personas, a usar armas para herir o a implicarse en peleas con golpes (Espinoza López & Tobar-Viera, 2023). El elemento distintivo consiste en que el sujeto usa partes de su propio cuerpo o un objeto externo para producir una lesión o daño (Buss & Perry, 1992). Buss (1961) la define como “un ataque contra un organismo por medio de partes del cuerpo (extremidades, dientes) o armas (cuchillo, garrote, pistola)” (pp. 4-5, traducción del autor).

La agresión verbal, por su parte, abarca aquellas conductas que se orientan a dañar a otros individuos mediante la crítica, el insulto, la burla, el sarcasmo y otras modalidades de expresión verbal (Espinoza López & Tobar-Viera, 2023). Buss (1961) la define con las siguientes palabras: “una respuesta vocal que envía estímulos nocivos a otro organismo. Los estímulos nocivos de la agresión física son el dolor y la lesión, mientras que los de la agresión verbal son el rechazo y la amenaza” (p. 6, traducción del autor).

La ira implica la activación fisiológica y la preparación para la agresión (Buss & Perry, 1992). Entendida como rasgo, se concibe como un temperamento general caracterizado por un bajo umbral de reactividad o como un patrón más estrecho de reactividad ante clases específicas de estímulos (Ramírez & Andreu, 2006). De acuerdo con estos autores, ello implica que los sentimientos de ira pueden presentarse como reacción ante múltiples detonantes comparativamente triviales (por ejemplo, una espera prolongada en la fila de un cajero), así como ante tipos particulares de situaciones (por ejemplo, la rivalidad, la exclusión o la percepción de recibir trato inequitativo).

Por último, la hostilidad consiste en sentimientos de mala voluntad e injusticia (Buss & Perry, 1992), y puede definirse como una evaluación negativa de personas y cosas que está acompañada frecuentemente por un deseo de dañarlas (Ramírez & Andreu, 2006). Espinoza López y Tobar-Viera (2023) la consideran como una desconfianza excesiva hacia las conductas o intenciones de otras personas. De acuerdo con Buss (1961), se define como “una respuesta verbal implícita que implica sentimientos negativos (mala voluntad) y evaluaciones negativas de personas y acontecimientos” (p. 12, traducción del autor).

En la actualidad, se han realizado múltiples estudios sobre las propiedades psicométricas del AQ en distintos idiomas, por ejemplo: alemán (von Collani & Werner, 2005), húngaro (Gerevich et al., 2007; Zimonyi et al., 2021) y japonés (Nakano, 2001). En el ámbito iberoamericano, se han analizado dichas propiedades en distintas poblaciones como alumnas peruanas de enfermería (Córdova-Delgado et al., 2025), adultos argentinos (García-Batista et al., 2022) y personas en privación de libertad en España (Cruz et al., 2024), entre otras.

Estudios de validación del AQ en población joven y adolescente se han realizado en España (Andreu Rodríguez et al., 2002), El Salvador (Sierra & Gutiérrez Quintanilla, 2007), Argentina (Reyna et al., 2011), Perú (Pomahuacre-Carhuayal et al., 2024), Puerto Rico (Cruz-Pérez et al., 2013), Chile (García-Fernández et al., 2015), Colombia (Agudelo Gallego et al., 2024; Javela et al., 2023) y Ecuador (Espinoza López & Tobar-Viera, 2023), entre otros. No obstante, hasta el momento no se ha identificado ningún estudio sobre las propiedades psicométricas de este instrumento en adolescentes mexicanos. Se encontró solamente un estudio realizado con una muestra de 213 personas de 13 a 51 años, distribuidas en siete estados de la república mexicana (Buendía-Lozada et al., 2019), pero ningún estudio enfocado específicamente en población adolescente.

Con base en estos antecedentes, la presente investigación se planteó como objetivo general evaluar la validez estructural y la confiabilidad del AQ en una muestra de adolescentes mexicanos de un plantel educativo de nivel medio superior. Para ello se plantearon los siguientes objetivos específicos: a) realizar un análisis factorial confirmatorio (AFC) del AQ para evaluar la validez de su estructura factorial; y b) evaluar la confiabilidad de sus escalas y del instrumento total, a través del estudio de su consistencia interna.

La pregunta principal de investigación fue la siguiente: ¿Qué evidencias de validez estructural y de confiabilidad presenta el AQ en una muestra de adolescentes mexicanos de nivel educativo medio superior?

Las hipótesis son las siguientes. H1: El modelo original de cuatro factores correlacionados (agresión física, agresión verbal, ira y hostilidad) del AQ presentará índices de bondad de ajuste adecuados. H2: se obtendrán coeficientes de consistencia interna aceptables tanto para el instrumento completo como para cada una de sus cuatro escalas.

Metodología

Tipo de estudio

La investigación se desarrolló como un estudio empírico de enfoque cuantitativo. Se adoptó un diseño transversal en el que la información se obtuvo a partir de una única aplicación del cuestionario a los participantes (Hernández Sampieri et al., 2010). Se trató de un estudio de tipo instrumental, orientado al análisis de las propiedades psicométricas de un instrumento de medición psicológica (Ato et al., 2013; Montero & León, 2002).

Participantes

Los participantes de la presente investigación fueron 683 estudiantes de un plantel de educación media superior de sostenimiento público ubicado en el área metropolitana de Guadalajara, Jalisco, México, de los cuales 52.9% eran mujeres y 47.1% eran hombres, con edades de 15 a 19 años ($M = 16.57$, $DE = 1.02$). El muestreo fue no probabilístico e intencional. Los criterios de inclusión fueron: ser estudiante de nivel medio superior, tener entre 15 y 19 años y encontrarse matriculado dentro del centro educativo. Los criterios de exclusión fueron: no ser estudiante de nivel medio superior, no tener entre 15 y 19 años cumplidos al momento del estudio, así como no encontrarse matriculado dentro del centro educativo.

Instrumento

Se utilizó el Cuestionario de Agresión (*The Aggression Questionnaire*, AQ) de Buss y Perry (1992), el cual está constituido por 29 ítems. La versión utilizada en el presente estudio fue la adaptación al idioma español realizada por Andreu Rodríguez et al. (2002), la cual se compone de los mismos 29 ítems del instrumento original, distribuidos en sus cuatro escalas: Agresión física (ítems 1, 5, 9, 13, 17, 21, 24 de puntuación inversa, 27 y 29), Agresión verbal (ítems 2, 6, 10, 14 y 18), Ira (ítems 3, 7, 11, 15 de puntuación inversa, 19, 22 y 25), y Hostilidad (ítems 4, 8, 12, 16, 20, 23, 26 y 28). Cada ítem utiliza un formato de respuesta tipo Likert de cinco opciones que va de 1 = “Completamente falso” hasta 5 = “Completamente verdadero”. Para la presente investigación el instrumento se aplicó conservando la redacción original de la versión de Andreu Rodríguez et al. (2002), sin someterlo a un proceso previo de adaptación lingüística o cultural. Esta decisión se basó en el supuesto de que la versión empleada utiliza un léxico estándar inteligible en función de la

competencia lectora esperada en estudiantes de nivel medio superior en el contexto local. Con ello, se priorizó mantener la fidelidad a la formulación original para preservar la comparabilidad directa con la literatura previa.

Procedimiento

Se obtuvo la autorización de la dirección del plantel educativo, cuyas autoridades realizaron las gestiones correspondientes con los padres, madres o tutores para la participación del alumnado de acuerdo con los procedimientos internos de la institución. Para la coordinación de los horarios de aplicación se contó con el apoyo del departamento de orientación educativa. Los cuestionarios se aplicaron en cada aula por el investigador, con la colaboración del profesorado.

El asentimiento del estudiantado se obtuvo mediante el siguiente procedimiento. Antes de iniciar la aplicación en cada grupo, se explicó al estudiantado el objetivo del estudio y las condiciones de participación, enfatizando su carácter voluntario y anónimo. Asimismo, se aclaró que no existían respuestas correctas o incorrectas y que la información sería analizada de manera estadística con fines exclusivamente científicos. Se verificó que estas condiciones fueran comprendidas por el alumnado, atendiendo sus preguntas y solicitudes de aclaración antes de iniciar la aplicación. La aplicación comenzó únicamente cuando el estudiantado manifestaba comprender la información proporcionada y su disposición a colaborar, lo cual constituyó su asentimiento para participar en el estudio. Quienes decidieron no participar pudieron elegir con entera libertad el permanecer en el aula o salir de ella, sin necesidad de justificar su decisión ni de comunicarla al investigador o al resto del grupo. Asimismo, durante la aplicación cualquier estudiante tuvo la opción de interrumpir su participación en el momento que lo decidiera, sin necesidad de justificar su decisión, así como retirarse o permanecer dentro del aula sin consecuencia alguna.

Análisis estadístico

Como fase previa al análisis principal, se realizó un análisis exploratorio de los datos. Se hizo una revisión y tratamiento de los datos perdidos, siguiendo las recomendaciones de Ferrando et al. (2022). La base de datos originalmente contenía 739 participantes. En primer lugar, se examinó si algún participante había dejado de responder más del 50% de los ítems del instrumento, caso en el cual se recomienda eliminar al participante del análisis debido a la falta de información suficiente para la estimación de sus puntuaciones. No se excluyó a

ningún participante bajo este criterio. En segundo lugar, se revisó si se presentaban ítems que tendieran a no ser respondidos por los participantes, ya que podría tratarse de ítems defectuosos o no adecuados para la población de estudio. Sin embargo, todos los ítems presentaron altas tasas de repuesta, ya que el ítem que presentó mayor número de casos perdidos fue el 16 con solamente nueve, lo cual representó solamente el 1.22% de casillas no respondidas.

Hecho lo anterior, se revisó que la presencia de datos perdidos no fuera superior al 5% de las respuestas totales al cuestionario, ya que en caso contrario debería pausarse el estudio para analizar la razón de la falta de respuestas. No obstante, este requisito se cumplió siendo solamente 0.56% de datos perdidos en total.

Los datos faltantes se trataron mediante imputación múltiple. De acuerdo con Ferrando et al. (2022), bajo los supuestos señalados, el procedimiento de imputación es preferible a la alternativa de eliminar los casos de la base de datos, ya que esto último podría homogeneizar la muestra de manera artificial. Antes de hacer la imputación se invirtieron las puntuaciones de los ítems 15 y 24, de acuerdo con la configuración original de la prueba (Andreu Rodríguez et al., 2002; Buss & Perry, 1992). Se hizo un procedimiento de imputación de los datos faltantes en cada una de las cuatro escalas del instrumento (Agresión física, Agresión verbal, Ira y Hostilidad), utilizando la opción de método automático la cual empleó el algoritmo de especificación condicional completa. Las variables incluidas en cada modelo fueron los ítems correspondientes a cada escala. Todos estos procedimientos se realizaron en SPSS v. 21 (IBM Corp., 2012).

Posteriormente se procedió a la detección de casos atípicos mediante el cálculo de la distancia de Mahalanobis (1936/2018). De acuerdo con Kline (2016), se deben excluir aquellos casos detectados que presenten valores de $p < .001$. Se detectaron 56 casos, los cuales fueron excluidos del análisis. Así, la base de datos que fue utilizada para el AFC quedó conformada por 683 participantes.

Hecho lo anterior, se realizó un examen del cumplimiento de supuestos para el análisis factorial. Se examinó la normalidad univariada de los ítems mediante estadística descriptiva, considerando valores de asimetría y curtosis como buenos en el intervalo de -1 a +1 y como aceptables de -2 a +2, según George y Mallery (2020). La normalidad multivariada se evaluó a través de la prueba de Mardia (1970) considerando aceptables valores no estandarizados de curtosis multivariada inferiores a 70 (Rodríguez Ayán & Ruiz Díaz, 2008).

La pertinencia del AFC se analizó a través de la prueba de Bartlett (1950), el índice de Kaiser-Meyer-Olkin (KMO) y la correspondiente medida de adecuación (*Measure of Sampling Adequacy*, MSA). Se espera que la prueba de esfericidad de Bartlett resulte estadísticamente significativa. Por su parte, de acuerdo con Ferrando et al. (2022), solo valores de KMO superiores a .75 son admisibles. En la medida de adecuación (MSA) se usa la misma formulación del KMO pero de manera individual para cada ítem. Valores de MSA menores de .50 se consideran inaceptables, y se recomienda eliminar esos ítems (Ferrando et al., 2022).

Para cumplir con el primer objetivo de esta investigación, se llevó a cabo un AFC del AQ utilizando el método de estimación de mínimos cuadrados ponderados con ajuste por media y varianza (*Weighted Least Squares Mean and Variance Adjusted*, WLSMV), debido a la naturaleza ordinal y a la distribución no normal de los datos (Rogers, 2024). El análisis se implementó en JASP v. 0.19.3 (JASP Team, 2025). Se hipotetizó un modelo de cuatro factores correlacionados en congruencia con la versión original del instrumento: Agresión física (9 ítems), Agresión verbal (5 ítems), Ira (7 ítems) y Hostilidad (8 ítems). Se estipularon cargas factoriales estandarizadas de $\geq .50$ como aceptables y de $\geq .70$ como ideales (Hair et al., 2019) para el modelo evaluado.

Para evaluar la bondad de ajuste se utilizó el cociente entre Chi-cuadrado y sus grados de libertad (χ^2/gl) también conocido como Chi-cuadrado normada, el índice de bondad de ajuste comparativo (*Comparative Fit Index*, CFI), el índice de Tucker-Lewis (TLI), el residuo cuadrático medio de aproximación (*Root Mean Square Error of Approximation*, RMSEA) y el residuo estandarizado cuadrático medio (*Standardized Root Mean Square Residual*, SRMR). Se consideraron los siguientes puntos de corte: $\chi^2/gl < 5$ (Hair et al., 1999), y CFI y TLI $\geq .95$, SRMR $< .08$, RMSEA $< .06$ (Hu & Bentler, 1999), aunque un RMSEA $< .08$ también ha sido considerado admisible (Schumacker & Lomax, 2016).

Con respecto al segundo objetivo particular del presente estudio, se utilizaron dos coeficientes de confiabilidad: alfa de Cronbach (1951) y omega de McDonald (2011). Valores $\geq .70$, .80 y .90 son considerados como aceptables, buenos y excelentes en ambos casos (Stefana et al., 2025). Se utilizaron los dos índices debido a que el omega se considera actualmente como una alternativa más precisa (Ventura-León & Caycho-Rodríguez, 2017); pero el coeficiente alfa ha sido utilizado en estudios psicométricos anteriores del AQ (Espinoza López & Tobar-Viera, 2023; Javela et al., 2023), por lo tanto, se incluye para facilitar su comparación con estudios previos. En ambos casos se consideró el uso de

intervalos de confianza del 95% debido a que actualmente se recomienda informarlos como indicadores que permiten valorar la precisión de la estimación de los coeficientes (Ventura-León, 2018).

Resultados

Examen de supuestos

En la Tabla 1 se presentan los resultados del análisis descriptivo de los ítems. Como se observa, doce de los ítems presentaron valores de asimetría o curtosis fuera del intervalo de -1 a +1, aunque todos se ubicaron dentro del intervalo de -2 a +2.

Tabla 1. Estadísticos descriptivos de los ítems del Cuestionario de Agresión (AQ)

Ítem	<i>M</i>	<i>DE</i>	Asimetría	Curtosis
1	2.00	1.16	0.86	-0.33
2	2.88	1.15	-0.19	-0.81
3	3.06	1.19	-0.28	-0.77
4	2.08	1.09	0.65	-0.59
5	2.58	1.42	0.33	-1.27
6	2.53	1.12	0.15	-0.78
7	2.72	1.24	0.12	-0.96
8	2.60	1.25	0.23	-1.00
9	2.96	1.38	-0.08	-1.23
10	2.81	1.20	-0.09	-1.01
11	2.65	1.40	0.25	-1.24
12	2.53	1.17	0.18	-0.92
13	1.99	1.10	0.83	-0.28
14	2.30	1.08	0.36	-0.76
15	2.86	1.07	0.23	-0.35
16	2.94	1.20	-0.12	-0.85
17	2.55	1.33	0.33	-1.05
18	2.22	1.22	0.73	-0.46
19	2.29	1.26	0.66	-0.64
20	2.58	1.29	0.26	-1.00
21	2.13	1.20	0.73	-0.52
22	2.11	1.14	0.64	-0.60
23	2.89	1.31	-0.05	-1.10
24	3.23	1.43	-0.18	-1.25
25	2.59	1.30	0.32	-0.99
26	2.75	1.29	0.07	-1.09
27	1.90	1.16	1.04	-0.01
28	3.12	1.17	-0.32	-0.67
29	2.63	1.36	0.26	-1.19

Nota. *N* = 683.

Fuente: elaboración propia

Por su parte, se obtuvieron valores estadísticamente significativos tanto de asimetría multivariada ($b_{1p} = 65.86$, $\chi^2_{(4495)} = 7496.58$, $p < .001$) como de curtosis multivariada ($b_{2p} = 984.55$; $z = 26.36$, $p < .001$). Como se aprecia, el valor del coeficiente de curtosis multivariada fue considerablemente superior al límite de 70, por lo que no se cumplió con el supuesto de normalidad multivariada; en consecuencia, la estimación se realizó con WLSMV de acuerdo con el procedimiento previsto en el apartado de Método.

Primer modelo: 29 ítems, cuatro factores correlacionados

Se llevó a cabo un primer análisis con los 29 ítems del instrumento. La prueba de esfericidad de Bartlett resultó estadísticamente significativa $\chi^2_{(406)} = 10090$, $p < .001$, y el índice KMO mostró un valor de .934. Asimismo, todos los ítems presentaron valores de medida de adecuación muestral (MSA) superiores a .50. Los ítems 15 y 24 presentaron los valores más bajos, siendo de .51 y .78, respectivamente.

Con el objetivo de evaluar la estructura factorial del instrumento, se llevó a cabo un AFC sobre los 29 ítems, utilizando el estimador WLSMV. El modelo especificado mostró un ajuste global de $\chi^2_{(371)} = 1795.01$, $p < .001$, así como los siguientes índices de bondad de ajuste: $\chi^2/gl = 4.84$, CFI = .915, TLI = .906, RMSEA = .075 (IC 90% [.072, .079]) y SRMR = .063.

En cuanto a los parámetros del modelo, la mayoría de los ítems presentó cargas factoriales estandarizadas estadísticamente significativas ($p < .001$) en el rango de $\lambda = .41$ a .81, con excepción de los ítems 15, 24 y 28. El ítem 15 mostró una carga factorial muy baja y no significativa estadísticamente ($\lambda = -.005$, $p = .900$). El ítem 24 presentó una carga estadísticamente significativa pero baja y negativa ($\lambda = -.262$, $p < .001$). El ítem 28 presentó también una carga factorial significativa pero de magnitud baja ($\lambda = .305$, $p < .001$).

Los coeficientes de determinación (R^2) indicaron que la varianza explicada por los factores latentes fue adecuada para la mayoría de los ítems en un intervalo de .26 a .66. Los ítems 2 y 3 obtuvieron valores de .166. Nuevamente presentaron problemas los ítems 15, 24 y 28 con coeficientes de determinación de .069, .0000219 y .093.

Con base en estos resultados del AFC, se decidió eliminar los ítems 15, 24 y 28, debido a su bajo desempeño psicométrico, en función de lo siguiente. Como se mencionó, el ítem 15 presentó una carga factorial estandarizada no significativa y una varianza explicada muy baja. Por su parte, el ítem 24 mostró una carga factorial baja y negativa, así como un coeficiente de determinación prácticamente nulo. Además, el ítem 28 presentó una carga

factorial estandarizada baja y un coeficiente de determinación reducido, lo que indica una contribución limitada al factor latente (hostilidad). Dado este bajo desempeño psicométrico, se estimó apropiado comparar este primer modelo con otro más parsimonioso que prescindiera de los tres ítems mencionados.

Segundo modelo: 26 ítems, cuatro factores correlacionados

Con base en los resultados anteriores, se llevó a cabo un segundo AFC con 26 ítems del instrumento (modelo revisado). La prueba de esfericidad de Bartlett resultó estadísticamente significativa, $\chi^2_{(325)} = 9534, p < .001$. El índice KMO alcanzó un valor global de .941, y todos los ítems presentaron valores de MSA elevados, en el intervalo de .899 a .972.

Como se indicó, el estimador fue WLSMV. El modelo especificado mostró un ajuste global de $\chi^2_{(293)} = 1359.14, p < .001$, así como los siguientes índices de bondad de ajuste: $\chi^2/gl = 4.64$, CFI = .933, TLI = .925, RMSEA = .073 (IC 90% [.069, .077]) y SRMR = .056.

En relación con los parámetros del modelo, como se puede observar en la Tabla 2, los ítems presentaron cargas factoriales estandarizadas en el rango de .41 a .81, lo que indica una adecuada relación entre los ítems y sus respectivos factores latentes.

Asimismo, los coeficientes de determinación (R^2) mostraron niveles de varianza explicada superiores a .50 en 21 ítems, equivalente al 80.77% del total. Dos ítems (4 y 14) presentaron coeficientes de determinación superiores a .30, mientras que solo los ítems 2 y 3 mostraron valores menores a .20.

Tabla 2. Resultados del análisis factorial confirmatorio del Cuestionario de Agresión (AQ), Modelo revisado

Ítem	Carga factorial estandarizada (λ)	EE	p	IC del 95%		R ²
				Inferior	Superior	
Factor 1. Agresión física						
1	0.749	0.025	< .001	0.699	0.799	0.561
5	0.745	0.021	< .001	0.703	0.786	0.555
9	0.690	0.024	< .001	0.644	0.737	0.476
13	0.771	0.022	< .001	0.727	0.816	0.595
17	0.710	0.023	< .001	0.665	0.755	0.504
21	0.814	0.019	< .001	0.777	0.852	0.663
27	0.766	0.022	< .001	0.722	0.810	0.587
29	0.638	0.027	< .001	0.585	0.691	0.407
Factor 2. Agresión verbal						
2	0.405	0.033	< .001	0.341	0.469	0.164
6	0.644	0.026	< .001	0.592	0.696	0.415
10	0.711	0.021	< .001	0.669	0.753	0.506
14	0.620	0.027	< .001	0.567	0.672	0.384
18	0.728	0.023	< .001	0.684	0.773	0.530
Factor 3. Ira						
3	0.409	0.034	< .001	0.343	0.474	0.167
7	0.639	0.024	< .001	0.591	0.687	0.408
11	0.746	0.021	< .001	0.705	0.786	0.556
19	0.788	0.019	< .001	0.751	0.824	0.620
22	0.804	0.018	< .001	0.768	0.839	0.646
25	0.647	0.024	< .001	0.600	0.695	0.419
Factor 4. Hostilidad						
4	0.550	0.034	< .001	0.482	0.617	0.302
8	0.636	0.028	< .001	0.582	0.690	0.404
12	0.685	0.028	< .001	0.630	0.740	0.470
16	0.644	0.027	< .001	0.591	0.698	0.415
20	0.643	0.028	< .001	0.588	0.697	0.413
23	0.503	0.034	< .001	0.437	0.569	0.253
26	0.714	0.025	< .001	0.665	0.762	0.509
Nota. Método de estimación: mínimos cuadrados ponderados con ajuste por media y varianza (WLSMV) implementado en JASP v. 0.19.3 (JASP Team, 2025). Modelo de						

Ítem	Carga factorial estandarizada (λ)	EE	p	IC del 95%		R ²
				Inferior	Superior	
cuatro factores correlacionados con 26 ítems. EE = error estándar. R ² = coeficientes de determinación estandarizados.						

Fuente: elaboración propia

Correlaciones entre factores

Las correlaciones entre los factores latentes resultaron ser directas y estadísticamente significativas, tal como puede apreciarse en la Tabla 3. Como se observa, la correlación más elevada se presentó entre agresión verbal e ira, en segundo lugar entre agresión física y verbal, y en tercer lugar entre agresión física e ira. La hostilidad presentó su correlación más alta con ira, seguida de la agresión verbal y finalmente con la agresión física.

Tabla 3. Correlaciones entre los factores latentes del Cuestionario de Agresión (AQ)

	Agresión física	Agresión verbal	Ira	Hostilidad
Agresión física	1	.903	.867	.654
Agresión verbal		1	.973	.792
Ira			1	.858
Hostilidad				1
<i>Nota.</i> Correlaciones interfactoriales estandarizadas estimadas a partir del análisis factorial confirmatorio. Todas las correlaciones fueron estadísticamente significativas ($p < .001$).				

Fuente: elaboración propia.

Consistencia interna

En la Tabla 4 se presentan los resultados del análisis de la consistencia interna del instrumento total y de cada una de sus escalas, estimada mediante alfa de Cronbach (α) y omega de McDonald (ω), tal como se señaló en el apartado de Método. Como se puede apreciar, todos los índices fueron superiores a .70, y en algunos casos superiores a .80. Para el instrumento total (26 ítems), los índices resultantes se ubicaron por encima de .90.

Tabla 4. Resultados del análisis de consistencia interna del Cuestionario de Agresión (AQ)

	ω	IC (95%)	α	IC (95%)
Agresión física	.868	[.851, .883]	.867	[.850, .882]
Agresión verbal	.727	[.694, .758]	.724	[.689, .755]
Ira	.801	[.775, .823]	.794	[.766, .816]
Hostilidad	.782	[.755, .806]	.777	[.750, .801]
Total	.928	[.920, .934]	.927	[.919, .933]
<i>Nota.</i> α = coeficiente alfa de Cronbach; ω = coeficiente omega de McDonald. Los intervalos de confianza (IC 95 %) se obtuvieron mediante <i>bootstrap</i> no paramétrico (1000 remuestras).				

Fuente: elaboración propia.

Discusión

El presente estudio partió de la siguiente pregunta: ¿Qué evidencias de validez estructural y de confiabilidad presenta el AQ en una muestra de adolescentes mexicanos de nivel educativo medio superior?

Como se pudo observar, el modelo revisado con 26 ítems y cuatro factores correlacionados mostró un ajuste parcialmente aceptable. El índice Chi-cuadrado normada (χ^2/gl) presentó un valor adecuado al ser menor a cinco (Hair et al., 1999). El SRMR cumplió con el criterio establecido al situarse por debajo de .06 (Hu & Bentler, 1999). Por su parte, el RMSEA no el alcanzó punto de corte estricto de $< .06$ (Hu & Bentler, 1999); sin embargo presentó un valor menor a .08 considerándose como admisible (Schumacker & Lomax, 2016). No obstante lo anterior, los valores de CFI y TLI no alcanzaron el criterio estipulado de .95 ya que se situaron por debajo de ese umbral (Hu & Bentler, 1999). En conjunto, estos resultados proporcionan evidencia favorable aunque no concluyente sobre la adecuación de la estructura factorial propuesta, por lo que su interpretación debe realizarse con cautela y considerando el conjunto de los índices de ajuste.

Los hallazgos de este estudio indican también que el modelo de 26 ítems constituye una solución más parsimoniosa en comparación con el modelo inicial de 29 ítems, observándose una mejora en todos los índices de bondad de ajuste considerados, lo que respalda la eliminación de los ítems que presentaron bajo desempeño psicométrico (ítems 15, 24 y 28). Los mencionados ítems eliminados presentan características que podrían explicar su bajo desempeño psicométrico. Los ítems 15 y el 24 se encuentran formulados en sentido inverso, una característica que frecuentemente se asocia con dificultades de interpretación que reducen su asociación con el constructo latente (Elek et al., 2025). Por su parte, el ítem 28 representa una manifestación específica de desconfianza interpersonal que podría no

reflejar de manera consistente el núcleo conceptual de la hostilidad en la presente muestra, pudiendo tratarse de un constructo relacionado pero diferenciable (Mo et al., 2025). Estas explicaciones se formulan de manera hipotética, y por lo tanto, deberán ser evaluadas en posteriores estudios. También se recomienda que futuras investigaciones analicen si los ítems eliminados requieren adaptación cultural o reformulación para lograr su recuperación y mejorar su funcionamiento.

Por otra parte, en el modelo revisado la mayor parte de los ítems presentó cargas factoriales aceptables y congruentes con la estructura teórica propuesta. En particular, 12 ítems presentaron cargas superiores a .50, mientras que otros 12 alcanzaron el valor de .70 considerado como ideal de acuerdo con los criterios estipulados (Hair et al., 2019). Únicamente los ítems 2 y 3 presentaron cargas menores a .50, aunque superiores a .40. En consecuencia, se decidió conservar estos ítems en el modelo final, si bien se recomienda continuar evaluando con atención su desempeño en estudios posteriores.

Los resultados del presente estudio coinciden parcialmente con los de Andreu Rodríguez et al. (2002) y Agudelo Gallego et al. (2024) realizados en España y Colombia, respectivamente, donde se obtuvieron hallazgos favorables para el modelo con cuatro factores, solo que en esos estudios los ítems 19, 23 y 29 fueron los que presentaron problemas de rendimiento psicométrico. Por otra parte, este trabajo coincidió con el de Buendia-Lozada et al. (2019), realizado en México con 213 personas de 13 a 51 años. En dicho estudio los autores encontraron que se replicaba el modelo original de cuatro factores con 29 ítems, y un índice de confiabilidad alfa de .91. Esta relativa convergencia de los resultados de los estudios mencionados con los obtenidos en el presente trabajo contribuye en cierta medida a dar soporte al modelo multidimensional de la agresión medida por el AQ, considerando cuatro factores: agresión física, agresión verbal, ira y hostilidad; no obstante, sugiere también la necesidad de seguir evaluando en trabajos posteriores tanto el modelo como el desempeño específico de los ítems.

Cabe mencionar que algunos estudios llevados a cabo en Latinoamérica no utilizaron la versión de 29 ítems de Buss y Perry (1992) traducida por Andreu Rodríguez et al. (2002), que se empleó en el presente trabajo. Por ejemplo, Castrillón et al. (2004) tomaron como punto de partida una versión de 40 ítems que, tras el proceso de validación, quedó conformada por 19 ítems. Esa versión de 40 ítems fue retomada por López del Pino et al. (2009), y más recientemente Javela et al. (2023) aplicaron la versión de 19 ítems. Por este

motivo, no resultó viable establecer una comparación directa entre los hallazgos del presente estudio y los mencionados.

Por otra parte, conviene destacar que las correlaciones interfactoriales estimadas en el AFC del presente estudio fueron elevadas y estadísticamente significativas, y el patrón de relaciones hallado resultó consistente parcialmente con el modelo original en el sentido de que se trató de correlaciones directas entre todas los factores (Andreu Rodríguez et al., 2002; Buss & Perry, 1992). No obstante lo anterior, las elevadas correlaciones observadas, particularmente entre agresión verbal e ira y entre agresión física y agresión verbal, indican que la diferenciación empírica entre estas dimensiones debe interpretarse con cautela. En este sentido, aunque los resultados son compatibles con la organización multidimensional del instrumento, no permiten afirmar de manera concluyente una diferenciación nítida entre sus factores. Por ello, se recomienda continuar examinando en estudios posteriores la estabilidad de esta estructura factorial en muestras y contextos diversos, comparando estructuras alternativas teóricamente justificadas.

Finalmente, con relación al análisis de la consistencia interna, los coeficientes alfa de Cronbach y omega de McDonald resultaron adecuados en las escalas de Agresión verbal y Hostilidad, buenos en Agresión física, y excelentes para el instrumento total (George & Mallery, 2020; Stefana et al., 2025). En cuanto a la escala de Ira esta obtuvo una consistencia interna de nivel adecuado según el coeficiente alfa, y bueno según el coeficiente omega. Estos resultados aportan evidencias iniciales de consistencia interna de cada una de las escalas del instrumento y de la puntuación total, lo cual coincide con estudios previos (Andreu Rodríguez et al., 2002; Buendia-Lozada et al., 2019; Espinoza López & Tobar-Viera, 2023).

Como limitaciones de esta investigación, cabe señalar el hecho de que el muestreo se realizó de manera no probabilística e intencional y en un solo plantel educativo público del área metropolitana de Guadalajara, lo cual no permite la generalización de los hallazgos ni el establecimiento de baremos poblacionales representativos. Por tal motivo, se recomienda que futuras investigaciones utilicen un muestreo probabilístico y una muestra suficientemente representativa de la población objetivo, a fin de avanzar en el conocimiento de las propiedades psicométricas del AQ.

Así mismo, este trabajo se centró en examinar las evidencias de validez estructural y confiabilidad del AQ, pero no se analizaron otras propiedades psicométricas como la invarianza de medición por sexo, edad y grado, la validez de criterio, la validez discriminante y la estabilidad test-retest. Por tal razón se recomienda que en futuros trabajos se avance en

el estudio de más propiedades psicométricas del instrumento, en población adolescente mexicana.

Finalmente, conviene mencionar que la versión del AQ utilizada en el presente estudio (Andreu Rodríguez et al., 2002) no fue sometida a un proceso de adaptación lingüística o cultural para la población adolescente mexicana. Si bien la mayor parte de los ítems del instrumento mantuvieron un comportamiento psicométrico aceptable, se sugiere que futuras investigaciones incorporen fases previas de análisis dirigidas a obtener información específica relacionada con la equivalencia cultural y semántica del instrumento en este grupo poblacional a fin de profundizar en los presentes hallazgos y obtener evidencia adicional de validez del instrumento. Tales investigaciones podrían incorporar aproximaciones cualitativas o mixtas que permitan examinar cómo los adolescentes mexicanos interpretan los contenidos de los ítems, en especial los que resultaron eliminados del modelo revisado en el presente estudio, con el propósito de identificar posibles variaciones culturales y contextuales.

Conclusiones

En suma, los hallazgos del presente estudio aportan evidencia parcial a favor de las hipótesis planteadas. En primer lugar, los índices de bondad de ajuste obtenidos para el modelo revisado de cuatro factores correlacionados con 26 ítems proporcionan evidencia parcialmente aceptable y, por lo tanto, no concluyente, sobre su estructura factorial en la muestra estudiada, por lo que su interpretación debe llevarse a cabo con cautela y considerando el conjunto de índices de bondad de ajuste. Así mismo, las elevadas correlaciones interfactoriales observadas no permiten afirmar de manera concluyente una diferenciación nítida entre las dimensiones del modelo, por lo que se requiere evaluación adicional. En segundo lugar, los coeficientes de consistencia interna hallados tanto para el instrumento total como para cada una de sus escalas aportan evidencias de confiabilidad del AQ en muestras con características similares a las del presente estudio.

Estos resultados permitieron atender los objetivos planteados; no obstante, debido al carácter no probabilístico del muestreo y a la procedencia de los participantes de un solo plantel educativo, los hallazgos no pueden generalizarse a la población adolescente mexicana de bachillerato en su conjunto. Por ello, se requieren estudios posteriores con muestras más amplias, probabilísticas y procedentes de diversos planteles y contextos educativos para

evaluar la estabilidad y replicabilidad de la estructura factorial y de los indicadores de confiabilidad del instrumento.

Futuras líneas de investigación

Se recomienda realizar estudios de adaptación lingüística y cultural que permitan obtener información sobre la equivalencia semántica y cultural del instrumento, considerando metodologías cualitativas o mixtas. También se plantea la necesidad de replicar el modelo revisado de 26 reactivos en una muestra independiente, así como en muestras probabilísticas multicéntricas procedentes de distintos contextos educativos y geográficos, con la finalidad de evaluar la estabilidad de la estructura factorial del cuestionario y ampliar la posibilidad de generalización de los hallazgos. Asimismo, se requiere que futuras investigaciones analicen otras fuentes de evidencia de validez y propiedades psicométricas del AQ en población adolescente mexicana; en particular, realizar estudios de invarianza factorial es necesario para sustentar la comparabilidad de puntuaciones entre grupos por sexo, edad y grado escolar. También se requieren evidencias de validez convergente, discriminante y de criterio, incluyendo estudios de validez discriminante entre los propios factores del AQ. Además, para profundizar en el análisis psicométrico, convendría también comparar la versión revisada de 26 ítems con otras versiones alternativas del AQ disponibles en la literatura. Finalmente, una vez que se lograra confirmar suficientemente la estructura, confiabilidad, invarianza y adecuación cultural del instrumento, sería recomendable la construcción de baremos que permitan establecer normas para la interpretación estandarizada de las puntuaciones en población adolescente mexicana.

Referencias

- Agudelo Gallego, C. M., Montoya Zuluaga, P. A., Cárdenas Niño, L., Pino Montoya, J. W., & Gómez-Tabares, A. S. (2024). Psychometric Properties of the Aggression Questionnaire (AQ) in a sample of Colombian children and adolescents. *Suma Psicológica*, 31(2), 34–51. <https://doi.org/10.14349/sumapsi.2024.v31.n2.5>
- Andreu Rodríguez, J. M., Peña Fernández, M. E., & Graña Gómez, J. L. (2002). Adaptación psicométrica de la versión española del Cuestionario de Agresión. *Psicothema*, 14(2), 476–482. <https://www.psicothema.com/pdf/751.pdf>
- Ato, M., López, J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038–1059. <https://doi.org/10.6018/analesps.29.3.178511>
- Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology*, 3(2), 77–85. <https://doi.org/10.1111/j.2044-8317.1950.tb00285.x>
- Buendia-Lozada, E. R. P., Alatríste-Flores, V. G., Islas-Guerra, S., López-Alonso, J., & Chávez-Erives, A. (2019). Validación del Cuestionario de Agresión de Buss y Perry (AQ) en actividad física y deporte. *Lecturas: Educación Física y Deportes*, 24(255), 13–31. <https://efdeportes.com/efdeportes/index.php/EFDeportes/article/view/1165>
- Buss, A. (1961). *The psychology of aggression*. John Wiley & Sons, Inc.
- Buss, A., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63(3), 452–459. <https://doi.org/10.1037/0022-3514.63.3.452>
- Castrillón M., D. A., Ortiz T., P. A., & Vieco G., 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(2), 49–61. <https://doi.org/10.17533/udea.rfnsp.561>
- Chester, D. S., Crowe, M. L., Hyatt, C. S., & Miller, J. D. (2024). The structure of aggressive personality. *Journal of Personality*, 92(5), 1375–1393. <https://doi.org/10.1111/jopy.12895>
- Córdova-Delgado, M., Guevara-Saravia, M. A., Muñoz-Aguirre, M., Puza-Mendoza, G. M., & Oyola-García, A. E. (2025). Consistencia del Cuestionario de Agresión de Buss y Perry en estudiantes universitarias peruanas de Enfermería. *Revista Información Científica*, 104, e4794. <https://doi.org/10.5281/zenodo.14837148>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>

- Cruz, A. R., Aizpurúa, E., Jimeno Jiménez, M. V., Zabala Baños, C., Barbosa, F., & Ricarte Trives, J. J. (2024). Propiedades psicométricas de las versiones larga y corta del cuestionario de agresión AQ en una muestra de presos españoles. *Revista Española de Sanidad Penitenciaria*, 26(1), 9–17. <https://doi.org/10.18176/resp.00080>
- Cruz-Pérez, J. P., Padilla-Cotto, L., & Sayers-Montalvo, S. K. (2013). Validación del Cuestionario de Agresión con una muestra de adolescentes puertorriqueños/as. *Revista Interamericana de Psicología*, 47(3), 365–372.
- Elek, D., Cígler, H., Grüning, D. J., & Jezek, S. (2025). Advancing the psychometrics of reverse-keyed items: Enriching cognitive theory by a logical and linguistic perspective. *Frontiers in Psychology*, 16, a1684612. <https://doi.org/10.3389/fpsyg.2025.1684612>
- Espinoza López, I., & Tobar-Viera, A. (2023). Propiedades psicométricas del Cuestionario de Agresión (AQ) en estudiantes universitarios ecuatorianos. *Ecos de la Academia*, 9(18), 9–27. <https://doi.org/10.53358/ecosacademia.v9i18.907>
- Ferrando, P. J., Lorenzo-Seva, U., Hernández-Dorado, A., & Muñiz, J. (2022). Decálogo para el Análisis Factorial de los Ítems de un test. *Psicothema*, 34(1), 7–17. <https://doi.org/10.7334/psicothema2021.456>
- García-Batista, Z. E., Moretti, L. S., Franco, P., Medrano, L. A., & Mustaca, A. E. (2022). Evidencias de validez y confiabilidad de las puntuaciones del Cuestionario de Agresividad (AQ) para adultos argentinos. *Revista Iberoamericana de Diagnóstico y Evaluación RIDEP*, 62(1), 17–28. <https://doi.org/10.21865/RIDEP62.1.02>
- García-Fernández, J. M., Lagos-San Martín, N., Díaz-Herrero, A., Inglés, C. J., & Torregrosa, M. S. (2015). Psychometric properties of the “aggressiveness questionnaire” in Chilean adolescents: comparison of different versions. *Behavioral Psychology*, 23(3), 489–505. <http://rua.ua.es/dspace/handle/10045/57931>
- George, D., & Mallery, P. (2020). *IBM SPSS statistics 26 step by step: A simple guide and reference* (16th ed.). Taylor & Francis.
- Gerevich, J., Bácskai, E., & Czobor, P. (2007). The generalizability of the Buss-Perry Aggression Questionnaire. *International Journal of Methods in Psychiatric Research*, 16(3), 124–136. <https://doi.org/10.1002/mpr.221>
- Hair, J., Anderson, R., Tatham, R., & Black, W. (1999). *Análisis multivariante* (E. Prentice & D. Cano, trads.). Prentice Hall Iberia.

- Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th. ed.). Cengage Learning EMEA.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2010). *Metodología de la investigación* (5^a. ed.). McGraw-Hil / Interamericana Editores.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- IBM Corp. (2012). *IBM SPSS Statistics for Windows* (Version 21.0) [Computer software]. IBM Corp.
- JASP Team. (2025). *JASP* (Version 0.19.3) [Computer software]. University of Amsterdam.
- Javela, J. J., Bahamón, M. J., Rivera, D., & González-Reyes, R. E. (2023). Analysis of the psychometric properties of the Buss-Perry Aggression Questionnaire (AQ) in Colombian adolescents. *Revista de Psicología Clínica con Niños y Adolescentes*, 10(3), 62–69. <https://doi.org/10.21134/rpcna.2023.10.3.8>
- Jiménez Velázquez, O. (2025). Educación para la Paz en México: programas para alcanzar los Objetivos de la Agenda 2030. *Revista Construyendo Paz Latinoamericana*, 11(23). <https://doi.org/10.35600/25008870.2026.23.0398>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling*. The Guilford Press.
- Krahé, B. (2021). *The social psychology of aggression* (3rd ed.). Routledge.
- López del Pino, C., Sánchez Burón, A., Fernández-Martín, M. P., & Rodríguez Tapia, L. T. (2009). Propiedades psicométricas del cuestionario AQ aplicado a población adolescente. *Edupsykhé. Revista de Psicología y Educación*, 8(1), 79–94. <https://doi.org/10.57087/edupsykhe.v8i1.3819>
- Mahalanobis, P. C. (2018). On the generalised distance in statistics. *Sankhya A*, 80(Suppl. 1), 1–7. <https://doi.org/10.1007/s13171-019-00164-5> (Trabajo original publicado en 1936).
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519–530. <https://doi.org/10.2307/2334770>
- McDonald, R. P. (2011). *Test theory: A unified treatment*. Routledge.
- Mo, J., Roorda, D. L., van der Ark, L. A., & Orobio de Castro, B. (2025). Is mistrust in early adolescence referent-specific? Looking at the validity of different mistrust referents

- in sixth Graders. *The Journal of Early Adolescence*, 45(6), 724–745. <https://doi.org/10.1177/02724316241273344>
- Montero, I., & León, O. (2002). Clasificación y descripción de las metodologías de investigación en Psicología. *International Journal of Clinical and Health Psychology*, 2(3), 503–508. http://www.aepc.es/ijchp/articulos_pdf/ijchp-53.pdf
- Nakano, K. (2001). Psychometric evaluation on the Japanese adaptation of the Aggression Questionnaire. *Behaviour Research and Therapy*, 39(7), 853–858. [https://doi.org/10.1016/s0005-7967\(00\)00057-7](https://doi.org/10.1016/s0005-7967(00)00057-7)
- Pomahuacre-Carhuayal, J. W., Flores-Ramírez, M. E., Sánchez-Aliaga, L. L., Pomahuacre-Carhuayal, L. A., & Alva-Mendoza, J. C. (2024). Cuestionario de Agresión de Buss y Perry, forma reducida (AQ-SF): evidencias psicométricas en adolescentes peruanos. *Revista de Investigación en Psicología*, 27(2), e27560. <https://doi.org/10.15381/rinvp.v27i2.27560>
- Ramírez, J. M., & Andreu, J. M. (2006). Aggression, and some related psychological constructs (anger, hostility, and impulsivity). Some comments from a research project. *Neuroscience and Biobehavioral Reviews*, 30(3), 276–291. <https://doi.org/10.1016/j.neubiorev.2005.04.015>
- Reyna, C., Lello Ivacevich, M. G., Sanchez, A., & 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. <https://doi.org/10.21500/20112084.775>
- Rodríguez Ayán, M. N., & Ruiz Díaz, M. A. (2008). Atenuación de la asimetría y de la curtosis de las puntuaciones observadas mediante transformaciones de variables: Incidencia sobre la estructura factorial. *Psicológica*, 29(2), 205–227.
- Rodríguez Molinero, L., & Imaz Roncero, C. (2020). Agresividad y conducta violenta en la adolescencia. *Adolescere. Revista de Formación Continuada de la Sociedad Española de Medicina de la Adolescencia*, 8(1), 62.e1-62.e9 <https://www.adolescere.es/agresividad-y-conducta-violenta-en-la-adolescencia/>
- Rogers, P. (2024). Best practices for your confirmatory factor analysis: A JASP and lavaan tutorial. *Behavior Research Methods*, 56, 6634–6654. <https://doi.org/10.3758/s13428-024-02375-7>
- Schumacker, R. E., & Lomax, R. G. (2016). *A beginner's guide to structural equation modeling* (4th ed.). Taylor & Francis.

- Sierra, J. C., & Gutiérrez Quintanilla, J. R. (2007). Validation of the Spanish version of the Buss- Perry Aggression Questionnaire in Salvadoran university students. *Psicología y Salud*, 17(1), 103–113.
<https://psicologiaysalud.uv.mx/index.php/psicysalud/article/view/745/1325>
- Stefana, A., Damiani, S., Granziol, U., Provenzani, U., Solmi, M., Youngstrom, E. A., & Fusar-Poli, P. (2025). Psychological, psychiatric, and behavioral sciences measurement scales: Best practice guidelines for their development and validation. *Frontiers in Psychology*, 15, Article 1494261.
<https://doi.org/10.3389/fpsyg.2024.1494261>
- Ventura-León, J. L. (2018). Intervalos de confianza para coeficiente Omega: Propuesta para el cálculo. *Adicciones*, 30(1), 77–78. <https://doi.org/10.20882/adicciones.962>
- Ventura-León, J. L., & Caycho-Rodríguez, T. (2017). El Coeficiente Omega: Un método alternativo para la estimación de la confiabilidad. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, 15(1), 625–627.
<https://www.redalyc.org/articulo.oa?id=77349627039>
- 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(7), 1631–1643. <https://doi.org/10.1016/j.paid.2004.09.027>
- Zimonyi, S., Kasos, K., Halmai, Z., Csirmaz, L., Stadler, H., Rózsa, S., Székely, A., & Kotyuk, E. (2021). Hungarian validation of the Buss–Perry Aggression Questionnaire. Is the short form more adequate? *Brain and Behavior*, 11(2), e02043.
<https://doi.org/10.1002/brb3.2043>