

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.

Date Received: January 2026

Date Accepted: July 2026

Introduction

The study of aggression in adolescents is relevant to understanding the processes by which dynamics of violence and conflict are produced and developed in this vital stage (Rodríguez Molinero & Imaz Roncero, 2020); therefore, the knowledge derived from this study is fundamental for the design of training, prevention and care strategies within the framework of education for peace, aimed at promoting a culture of peace based on respect for human rights and the Sustainable Development Goals (Jiménez Velázquez, 2025).

Furthermore, a deeper understanding of aggression is relevant due to its close relationship with multiple key aspects of adolescent development (Pomahuacre-Carhuayal et al., 2024). According to Javela et al. (2023), adolescence is a stage marked by brain transformations that affect emotional regulation and decision-making, a stage in which aggressive behaviors have been associated with clinical problems, the development of antisocial trajectories, and consequences such as school dropout and unemployment. Therefore, the accurate identification of aggression could be relevant for the design of effective interventions.

In this context, it is essential to have psychometric instruments with evidence of validity and adequate reliability estimates for the study of aggression.

The *Aggression Questionnaire* (AQ) by Buss and Perry (1992) is probably the most widely known instrument worldwide for measuring dispositional aggression through self-reporting (Krahé, 2021). According to Gerevich et al. (2007), once published in 1992, the AQ quickly became a gold standard for measuring aggression. One of the advantages of this instrument is that it consists of several scales (hostility, anger, physical aggression, and verbal aggression), which allows for a multidimensional estimation of the degree of dispositional aggression exhibited by individuals. According to Chester et al. (2024), this perspective, which conceives of aggression as being comprised of behavioral, affective, and cognitive components, is currently the most widely adopted approach for studying aggression as a personality trait or characteristic.

In this model proposed by Buss and Perry (1992), aggression, understood as a personality trait, is considered a multidimensional construct comprised of four components. Physical and verbal aggression constitute the behavioral component, anger the affective component, and hostility the cognitive component.

Physical aggression refers to a propensity to strike other people, to use weapons to injure, or to engage in fistfights (Espinoza López & Tobar-Viera, 2023). The distinguishing element is that the subject uses parts of their own body or an external object to produce an injury or harm (Buss & Perry, 1992). Buss (1961) defines it as “an assault against an organism by means of body parts (limb, teeth) or weapons (knife, club, gun)” (pp. 4-5).

Verbal aggression, on the other hand, encompasses behaviors aimed at harming other individuals through criticism, insults, mockery, sarcasm, and other forms of verbal expression (Espinoza López & Tobar-Viera, 2023). Buss (1961) defines it as follows: “a vocal response that delivers noxious stimuli to another organism. The noxious stimuli delivered in physical aggression are pain and injury; the noxious stimuli delivered in verbal aggression are rejection and threat” (p. 6).

Anger involves physiological arousal and preparation for aggression (Buss & Perry, 1992). Understood as a trait, it is conceived as a general temperament characterized by a low threshold of reactivity or as a narrower pattern of reactivity to specific classes of stimuli (Ramírez & Andreu, 2006). According to these authors, this implies that feelings of anger can arise as a reaction to multiple comparatively trivial triggers (e.g., a prolonged wait in line at a checkout), as well as to particular types of situations (e.g., rivalry, exclusion, or the perception of receiving unfair treatment).

Finally, hostility consists of feelings of ill will and injustice (Buss & Perry, 1992) and can be defined as a negative evaluation of people and things that is frequently accompanied by a desire to harm them (Ramírez & Andreu, 2006). Espinoza López and Tobar-Viera (2023) consider it to be an excessive distrust of the behaviors or intentions of other people. According to Buss (1961), it is defined as “an implicit verbal response involving negative feelings (ill will) and negative evaluations of people and events” (p. 12).

Currently, multiple studies have been conducted on the psychometric properties of the AQ in different languages, for example: German (von Collani & Werner, 2005), Hungarian (Gerevich et al., 2007; Zimonyi et al., 2021) and Japanese (Nakano, 2001). In the Ibero-American context, these properties have been analyzed in different populations such as Peruvian nursing students (Córdova-Delgado et al., 2025), Argentine adults (García-Batista et al., 2022) and people deprived of liberty in Spain (Cruz et al., 2024), among others.

Validation studies of the AQ in young and adolescent populations have been conducted in Spain (Andreu Rodríguez et al., 2002), El Salvador (Sierra & Gutiérrez Quintanilla, 2007), Argentina (Reyna et al., 2011), Peru (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), and Ecuador (Espinoza López & Tobar-Viera, 2023), among others. However, to date, no studies have been identified on the psychometric properties of this instrument in Mexican adolescents. Only one study was found conducted with a sample of 213 people aged 13 to 51, distributed in seven states of the Mexican republic (Buendia-Lozada et al., 2019), but no study focused specifically on the adolescent population.

Based on this background, the present research aimed to evaluate the structural validity and reliability of the AQ in a sample of Mexican adolescents from a high school. To this end, the following specific objectives were established: a) to conduct a confirmatory factor analysis (CFA) of the AQ to evaluate the validity of its factor structure; and b) to evaluate the reliability of its scales and the overall instrument through the study of its internal consistency.

The main research question was the following: What evidence of structural validity and reliability does the AQ present in a sample of Mexican adolescents at the upper secondary education level?

The hypotheses are as follows. H1: The original four-factor model of the AQ (physical aggression, verbal aggression, anger, and hostility) will show adequate goodness-of-fit indices. H2: Acceptable internal consistency coefficients will be obtained for both the complete instrument and each of its four scales.

Methodology

Type of study

The research was conducted as an empirical study with a quantitative approach. A cross-sectional design was adopted, in which the information was obtained from a single application of the questionnaire to the participants (Hernández Sampieri et al., 2010). It was an instrumental study, oriented towards the analysis of the psychometric properties of a psychological measurement instrument (Ato et al., 2013; Montero & León, 2002).

Participants

The participants in this study were 683 students from a public upper secondary school located in the metropolitan area of Guadalajara, Jalisco, Mexico. Of these, 52.9% were female and 47.1% were male, with ages ranging from 15 to 19 years ($M = 16.57$, $SD = 1.02$). The sampling method was non-probabilistic and purposive. Inclusion criteria were: being an upper secondary student, being between 15 and 19 years old, and being enrolled at the school. Exclusion criteria were: not being an upper secondary student, not being between 15 and 19 years old at the time of the study, and not being enrolled at the school.

Instrument

The 29-item *Aggression Questionnaire* (AQ) developed by Buss and Perry (1992) was used. The version used in the present study was the Spanish language adaptation by Andreu Rodríguez et al. (2002), which consists of the same 29 items as the original instrument, distributed in its four scales: Physical Aggression (items 1, 5, 9, 13, 17, 21, 24 reverse-scored, 27 and 29), Verbal Aggression (items 2, 6, 10, 14 and 18), Anger (items 3, 7, 11, 15 reverse-scored, 19, 22 and 25), and Hostility (items 4, 8, 12, 16, 20, 23, 26 and 28). Each item uses a five-point Likert-type response format ranging from 1 = “Completely false” to 5 = “Completely true.” For this research, the instrument was applied using the original wording from the version by Andreu Rodríguez et al. (2002), without undergoing any prior linguistic or cultural adaptation. This decision was based on the assumption that the version used employs a standard vocabulary intelligible to the expected reading level of upper secondary school students in the local context. Therefore, maintaining fidelity to the original formulation was prioritized to preserve direct comparability with previous literature.

Procedure

Authorization was obtained from the school administration, whose authorities coordinated the necessary arrangements with parents or legal guardians to ensure student participation, in accordance with the institution's internal procedures. The educational guidance department assisted in coordinating the administration schedule. The questionnaires were administered in each classroom by the researcher, with the collaboration of the teachers.

Student assent was obtained through the following procedure. Before beginning the survey in each group, the students were informed of the study's objective and the conditions of participation, with emphasis on its voluntary and anonymous nature. It was also clarified

that there were no right or wrong answers and that the information would be analyzed statistically for purely scientific purposes. Students' understanding of these conditions was verified by addressing their questions and requests for clarification before the survey began. The survey commenced only when students indicated they understood the information provided and were willing to participate, which constituted their assent to participate in the study. Those who chose not to participate were free to remain in the classroom or leave, without needing to justify their decision or inform the researcher or the rest of the group. Similarly, during the survey, any student had the option to interrupt their participation at any time, without needing to justify their decision, and to leave or remain in the classroom without any consequences.

Statistical analysis

As a preliminary step to the main analysis, an exploratory analysis of the data was conducted. A review and processing of missing data was performed, following the recommendations of Ferrando et al. (2022). The database originally contained 739 participants. First, it was examined whether any participant had failed to answer more than 50% of the instrument items. In such cases, it was recommended that the participant be excluded from the analysis due to insufficient information for estimating their scores. No participants were excluded under this criterion. Second, it was reviewed whether there were items that tended to go unanswered by participants, as these could be defective or unsuitable for the study population. However, all items showed high response rates, with item 16 having the highest number of missing items at only nine, representing just 1.22% of unanswered items.

Having done the above, it was verified that the presence of missing data did not exceed 5% of the total questionnaire responses, since otherwise the study would have to be paused to analyze the reason for the missing responses. However, this requirement was met, with only 0.56% of the data missing in total.

The missing data were addressed using multiple imputation. According to Ferrando et al. (2022), under the stated assumptions, the imputation procedure is preferable to removing cases from the database, as the latter could artificially homogenize the sample. Before imputation, the scores for items 15 and 24 were reversed, according to the original test configuration (Andreu Rodríguez et al., 2002; Buss & Perry, 1992). An imputation procedure was performed for the missing data on each of the instrument's four scales

(Physical Aggression, Verbal Aggression, Anger, and Hostility), using the automatic method option, which employed the complete conditional specification algorithm. The variables included in each model were the items corresponding to each scale. All these procedures were performed using SPSS version 21 (IBM Corp., 2012).

Outliers were then detected using the Mahalanobis distance (1936/2018). According to Kline (2016), cases with *p-values* < .001 should be excluded. Fifty-six cases were detected and excluded from the analysis. Thus, the database used for the CFA consisted of 683 participants.

Having done the above, an examination of the fulfillment of assumptions for factor analysis was carried out. Univariate normality of the items was examined using descriptive statistics, considering skewness and kurtosis values as good in the range of -1 to +1 and as acceptable from -2 to +2, according to George and Mallery (2020). Multivariate normality was evaluated using Mardia's test (1970), considering unstandardized multivariate kurtosis values below 70 as acceptable (Rodríguez Ayán & Ruiz Díaz, 2008).

Data suitability for CFA was analyzed through the Bartlett test (1950), the Kaiser-Meyer-Olkin (KMO) index and the corresponding measure of sampling adequacy (MSA). Bartlett's test of sphericity is expected to be statistically significant. According to Ferrando et al. (2022), only KMO values above .75 are acceptable. The MSA uses the same KMO formulation but is calculated individually for each item. MSA values below .50 are considered unacceptable, and it is recommended that these items be removed (Ferrando et al., 2022).

To meet the first objective of this research, a CFA of the AQ was carried out using the weighted least squares estimation method with adjustment for mean and variance (WLSMV), due to the ordinal nature and non-normal distribution of the data (Rogers, 2024). The analysis was implemented in JASP v. 0.19.3 (JASP Team, 2025). A four-factor model was hypothesized, consistent with the original version of the instrument: Physical Aggression (9 items), Verbal Aggression (5 items), Anger (7 items), and Hostility (8 items). Standardized factor loadings of $\geq .50$ were stipulated as acceptable and $\geq .70$ as ideal (Hair et al., 2019) for the evaluated model.

To assess the goodness of fit, the ratio between chi-square and its degrees of freedom (χ^2/df), also known as the normed Chi-square, was used, as well as the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR). The following cut-off points were

considered: $\chi^2/df < 5$ (Hair et al., 1999), and CFI and TLI $\geq .95$, SRMR $< .08$, RMSEA $< .06$ (Hu & Bentler, 1999), although an RMSEA $< .08$ has also been considered acceptable (Schumacker & Lomax, 2016).

Regarding the second specific objective of this study, two reliability coefficients were used: Cronbach's alpha (1951) and McDonald's omega (2011). Values $\geq .70$, $.80$, and $.90$ are considered acceptable, good, and excellent, respectively, in both cases (Stefana et al., 2025). Both indices were used because omega is currently considered a more precise alternative (Ventura-León & Caycho-Rodríguez, 2017); however, the alpha coefficient has been used in previous psychometric studies of the AQ (Espinoza López & Tobar-Viera, 2023; Javela et al., 2023) and is therefore included to facilitate comparison with previous studies. In both cases, 95% confidence intervals were used because they are currently recommended as indicators that allow for assessing the precision of the coefficient estimates (Ventura-León, 2018).

Results

Assessment of assumptions

Table 1 presents the results of the descriptive analysis of the items. As can be seen, twelve of the items showed skewness or kurtosis values outside the range of -1 to $+1$, although all were located within the range of -2 to $+2$.

Table 1. Descriptive statistics of the items of the Aggression Questionnaire (AQ)

Item	<i>M</i>	<i>SD</i>	Skewness	Kurtosis
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
Note. <i>N</i> = 683.				

Source: own elaboration

For their part, statistically significant values were obtained for both multivariate skewness ($b_{1p} = 65.86$, $\chi^2_{(4495)} = 7496.58$, $p < .001$) as multivariate kurtosis ($b_{2p} = 984.55$; $z = 26.36$, $p < .001$). As can be seen, the value of the multivariate kurtosis coefficient was considerably higher than the limit of 70, so the assumption of multivariate normality was not met; consequently, the estimation was carried out with WLSMV in accordance with the procedure provided in the Method section.

First model: 29 items, four correlated factors

A preliminary analysis was conducted using the 29 items of the instrument. Bartlett's test of sphericity was statistically significant ($\chi^2_{(406)} = 10090, p < .001$), and the KMO index showed a value of .934. Likewise, all items presented MSA values greater than .50. Items 15 and 24 presented the lowest MSA values, at .51 and .78, respectively.

To evaluate the instrument's factor structure, a confirmatory factor analysis (CFA) was performed on the 29 items using the WLSMV estimator. The specified model showed an overall fit of $\chi^2_{(371)} = 1795.01, p < .001$, as well as the following goodness of fit indices: $\chi^2/df = 4.84$, CFI = .915, TLI = .906, RMSEA = .075 (90% CI [.072, .079]) and SRMR = .063.

Regarding the model parameters, most items showed statistically significant standardized factor loadings ($p < .001$) in the range of $\lambda = .41$ to .81, with the exception of items 15, 24, and 28. Item 15 showed a very low and statistically insignificant factor loading ($\lambda = -.005, p = .900$). Item 24 showed a statistically significant but low and negative loading ($\lambda = -.262, p < .001$). Item 28 also showed a significant but low factor loading ($\lambda = .305, p < .001$).

The coefficients of determination (R^2) indicated that the variance explained by the latent factors was adequate for most items in a range of .26 to .66. Items 2 and 3 obtained values of .166. Items 15, 24 and 28 again presented problems with coefficients of determination of .069, .0000219 and .093.

Based on these CFA results, items 15, 24, and 28 were removed due to their poor psychometric performance, for the following reasons. As mentioned, item 15 showed a non-significant standardized factor loading and very low explained variance. Item 24 showed a low and negative factor loading, as well as a practically zero coefficient of determination. Furthermore, item 28 showed a low standardized factor loading and a reduced coefficient of determination, indicating a limited contribution to the latent factor (hostility). Given this poor psychometric performance, it was deemed appropriate to compare this first model with a more parsimonious one that omitted the three aforementioned items.

Second model: 26 items, four correlated factors

Based on the previous results, a second CFA was conducted with 26 items from the instrument (revised model). Bartlett's test of sphericity was statistically significant ($\chi^2_{(325)} = 9534, p < .001$). The KMO index reached an overall value of .941, and all items presented high MSA values, in the range of .899 to .972.

As indicated, the estimator was WLSMV. The specified model showed an overall fit of $\chi^2_{(293)} = 1359.14$, $p < .001$, as well as the following goodness of fit indices: $\chi^2/df = 4.64$, CFI = .933, TLI = .925, RMSEA = .073 (90% CI [.069, .077]) and SRMR = .056.

Regarding the model parameters, as can be seen in Table 2, the items presented standardized factor loadings in the range of .41 to .81, which indicates an adequate relationship between the items and their respective latent factors.

Furthermore, the coefficients of determination (R^2) showed levels of explained variance greater than .50 in 21 items, equivalent to 80.77% of the total. Two items (4 and 14) presented coefficients of determination greater than .30, while only items 2 and 3 showed values less than .20.

Table 2. Results of the confirmatory factor analysis of the Aggression Questionnaire (AQ),
Revised Model

Item	Standardized factor loading (λ)	SE	p	95% CI		R^2
				Lower	Superior	
Factor 1. Physical aggression						
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. Verbal aggression						
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. Anger						
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. Hostility						
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
Note. Estimation method: Weighted least squares adjusted for mean and variance (WLSMV) implemented in JASP v. 0.19.3 (JASP Team, 2025). Four-factor correlated model with 26 items. SE = standard error. R^2 = standardized coefficients of determination.						

Source: own elaboration

Correlations between factors

The correlations between the latent factors were direct and statistically significant, as shown in Table 3. As can be seen, the highest correlation was between verbal aggression and anger, followed by physical and verbal aggression, and then physical aggression and anger. Hostility showed its highest correlation with anger, followed by verbal aggression, and finally physical aggression.

Table 3. Correlations between the latent factors of the Aggression Questionnaire (AQ)

	Physical aggression	Verbal aggression	Anger	Hostility
Physical aggression	1	.903	.867	.654
Verbal aggression		1	.973	.792
Anger			1	.858
Hostility				1
<i>Note.</i> Standardized interfactorial correlations estimated from confirmatory factor analysis. All correlations were statistically significant ($p < .001$).				

Source: own elaboration.

Internal consistency

Table 4 presents the results of the internal consistency analysis of the overall instrument and each of its scales, estimated using Cronbach's alpha (α) and McDonald's omega (ω), as described in the Methods section. As can be seen, all indices were greater than .70, and in some cases greater than .80. For the overall instrument (26 items), the resulting indices were above .90.

Table 4. Results of the internal consistency analysis of the Aggression Questionnaire (AQ)

	ω	CI (95%)	α	CI (95%)
Physical aggression	.868	[.851, .883]	.867	[.850, .882]
Verbal aggression	.727	[.694, .758]	.724	[.689, .755]
Anger	.801	[.775, .823]	.794	[.766, .816]
Hostility	.782	[.755, .806]	.777	[.750, .801]
Total	.928	[.920, .934]	.927	[.919, .933]
<i>Note.</i> α = Cronbach's alpha coefficient; ω = McDonald's omega coefficient. Confidence intervals (95% CI) were obtained using non-parametric <i>bootstrap</i> (1000 resamples).				

Source: own elaboration.

Discussion

The present study started from the following question: What evidence of structural validity and reliability does the AQ present in a sample of Mexican adolescents at the upper secondary education level?

As observed, the revised model with 26 items and four correlated factors showed a partially acceptable fit. The normed chi-square index (χ^2/df) was adequate, being less than five (Hair et al., 1999). The SRMR met the established criterion, falling below .06 (Hu & Bentler, 1999). The RMSEA, however, did not reach the strict cutoff point of $< .06$ (Hu & Bentler, 1999); however, it was less than .08, which is considered acceptable (Schumacker & Lomax, 2016). Nevertheless, the CFI and TLI values did not reach the stipulated criterion of .95, as they fell below that threshold (Hu & Bentler, 1999). Taken together, these results provide favorable but not conclusive evidence on the suitability of the proposed factor structure, so their interpretation should be done with caution and considering all the fit indices.

The findings of this study also indicate that the 26-item model is a more parsimonious solution compared to the initial 29-item model, showing an improvement in all the goodness-of-fit indices considered. This supports the elimination of the items that performed poorly (items 15, 24, and 28). These items have characteristics that could explain their poor psychometric performance. Items 15 and 24 are worded in reverse, a characteristic frequently associated with interpretation difficulties that reduce their association with the underlying construct (Elek et al., 2025). Item 28, on the other hand, represents a specific manifestation of interpersonal mistrust that may not consistently reflect the core concept of hostility in the present sample. It could be a related but distinct construct (Mo et al., 2025). These

explanations are hypothetical and, therefore, should be evaluated in subsequent studies. It is also recommended that future research analyze whether the removed items require cultural adaptation or reformulation to achieve their recovery and improve their functioning.

Furthermore, in the revised model, most items showed acceptable factor loadings consistent with the proposed theoretical structure. Specifically, 12 items had loadings greater than .50, while another 12 reached the .70 value considered ideal according to the stipulated criteria (Hair et al., 2019). Only items 2 and 3 had loadings lower than .50, although still higher than .40. Consequently, it was decided to retain these items in the final model, although it is recommended that their performance be carefully evaluated in subsequent studies.

The results of this study partially coincide with those of Andreu Rodríguez et al. (2002) and Agudelo Gallego et al. (2024), conducted in Spain and Colombia, respectively, where findings favorable to the four-factor model were obtained. However, in those studies, items 19, 23, and 29 were the ones that presented psychometric performance problems. Furthermore, this work coincided with that of Buendia-Lozada et al. (2019), conducted in Mexico with 213 people aged 13 to 51. In that study, the authors found that the original four-factor model with 29 items was replicated, with an alpha reliability index of .91. This relative convergence of the results of the aforementioned studies with those obtained in the present work contributes to some extent to supporting the multidimensional model of aggression measured by the AQ, considering four factors: physical aggression, verbal aggression, anger, and hostility. However, it also suggests the need to continue evaluating both the model and the specific performance of the items in subsequent work.

It is worth mentioning that some studies conducted in Latin America did not use the 29-item version of the Buss and Perry (1992) questionnaire, translated by Andreu Rodríguez et al. (2002), which was used in this study. For example, Castrillón et al. (2004) started with a 40-item version, which, after the validation process, was reduced to 19 items. This 40-item version was then used by López del Pino et al. (2009), and more recently, Javela et al. (2023) applied the 19-item version. Therefore, it was not feasible to establish a direct comparison between the findings of this study and those mentioned above.

Furthermore, it is worth noting that the interfactorial correlations estimated in the confirmatory factor analysis (CFA) of this study were high and statistically significant, and the pattern of relationships found was partially consistent with the original model in that there were direct correlations between all factors (Andreu Rodríguez et al., 2002; Buss & Perry,

1992). However, the high correlations observed, particularly between verbal aggression and anger, and between physical aggression and verbal aggression, indicate that the empirical differentiation between these dimensions should be interpreted with caution. In this sense, although the results are compatible with the multidimensional organization of the instrument, they do not allow for a conclusive assertion of a clear differentiation between its factors. Therefore, it is recommended that future studies continue to examine the stability of this factor structure in diverse samples and contexts, comparing it with theoretically justified alternative structures.

Finally, regarding the internal consistency analysis, Cronbach's alpha and McDonald's omega coefficients were adequate for the Verbal Aggression and Hostility scales, good for Physical Aggression, and excellent for the overall instrument (George & Mallery, 2020; Stefana et al., 2025). The Anger scale showed adequate internal consistency according to the alpha coefficient and good internal consistency according to the omega coefficient. These results provide initial evidence of internal consistency for each of the instrument's scales and the total score, which is consistent with previous studies (Andreu Rodríguez et al., 2002; Buendia-Lozada et al., 2019; Espinoza López & Tobar-Viera, 2023).

Limitations of this research include the fact that the sampling was non-probabilistic and purposive, and only at a single public school in the Guadalajara metropolitan area. This prevents the generalizability of the findings and the establishment of representative population norms. Therefore, it is recommended that future research use probabilistic sampling and a sufficiently representative sample of the target population to further our understanding of the psychometric properties of the AQ.

This study focused on examining the evidence for the structural validity and reliability of the AQ, but did not analyze other psychometric properties such as measurement invariance by sex, age, and grade, criterion validity, discriminant validity, and test-retest reliability. Therefore, it is recommended that future studies further investigate more of the instrument's psychometric properties in a Mexican adolescent population.

Finally, it is worth mentioning that the version of the AQ used in this study (Andreu Rodríguez et al., 2002) was not subjected to a linguistic or cultural adaptation process for the Mexican adolescent population. While most of the instrument's items maintained acceptable psychometric performance, it is suggested that future research incorporate preliminary analysis phases aimed at obtaining specific information related to the cultural and semantic equivalence of the instrument in this population group in order to delve deeper into the

present findings and obtain additional evidence of the instrument's validity. Such research could incorporate qualitative or mixed-methods approaches that allow for examining how Mexican adolescents interpret the content of the items, especially those that were eliminated from the revised model in this study, with the purpose of identifying possible cultural and contextual variations.

Conclusions

In summary, the findings of this study provide partial evidence in support of the hypotheses. First, the goodness-of-fit indices obtained for the revised four-factor model with 26 correlated items provide partially acceptable, and therefore inconclusive, evidence regarding its factor structure in the studied sample. Therefore, their interpretation should be carried out with caution and considering the entire set of goodness-of-fit indices. Likewise, the high interfactor correlations observed do not allow for a conclusive assertion of a clear differentiation between the dimensions of the model, thus requiring further evaluation. Second, the internal consistency coefficients found for both the overall instrument and each of its scales provide evidence of the reliability of the AQ in samples with characteristics similar to those of the present study.

These results allowed us to meet the stated objectives; however, due to the non-probabilistic nature of the sampling and the fact that the participants came from a single educational institution, the findings cannot be generalized to the Mexican high school population as a whole. Therefore, further studies with larger, probabilistic samples from diverse schools and educational contexts are needed to evaluate the stability and replicability of the factor structure and the instrument's reliability indicators.

Future lines of research

Linguistic and cultural adaptation studies are recommended to obtain information on the semantic and cultural equivalence of the instrument, considering qualitative or mixed methodologies. It is also recommended that the revised 26-item model be replicated in an independent sample, as well as in multicenter probability samples from diverse educational and geographic contexts, to evaluate the stability of the questionnaire's factor structure and broaden the generalizability of the findings. Furthermore, future research should analyze other sources of evidence for the validity and psychometric properties of the AQ in Mexican adolescents; in particular, factor invariance studies are necessary to support the comparability

of scores between groups by sex, age, and grade level. Evidence of convergent, discriminant, and criterion validity is also required, including discriminant validity studies among the AQ's own factors. In addition, to further the psychometric analysis, it would be advisable to compare the revised 26-item version with other alternative versions of the AQ available in the literature. Finally, once the structure, reliability, invariance, and cultural appropriateness of the instrument have been sufficiently confirmed, it would be advisable to construct norms that allow the establishment of standards for the standardized interpretation of scores in the Mexican adolescent population.

References

- 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., Alatraste-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/edupsykhé.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>