International Journal of Methods in Psychiatric Research
*Int. J. Methods Psychiatr. Res.* 16(3): 124–136 (2007)
Published online in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/mpr.221

# The generalizability of the Buss–Perry Aggression Questionnaire

JÓZSEF GEREVICH,[1,2] ERIKA BÁCSKAI,[1] PÁL CZOBOR[3,4]

1 Addiction Research Institute, Budapest, Hungary
2 ELTE University, Faculty of Orthopedagogics, Budapest, Hungary
3 Semmelweis University, Department of Psychiatry, Budapest, Hungary
4 Nathan Kline Institute for Psychiatric Research, Orangeburg, New York, USA

**Abstract**

*Aggressive and hostile behaviours and anger constitute an important problem across cultures. The Buss–Perry Aggression Questionnaire (AQ), a self-rating scale was published in 1992, and has quickly become the gold-standard for the measurement of aggression. The AQ scale has been validated extensively, but the validation focused on various narrowly selected populations, typically, on samples of college students. Individuals, however, who are at risk of displaying aggressive and hostile behaviours may come from a more general population. Therefore, it is important to investigate the scale's properties in such a population. The objective of this study was to examine the factorial structure and the psychometric properties of the AQ scale in a nationally representative sample of the Hungarian adult population.*

*A representative sample of 1200 subjects was selected by a two-step procedure. The dimensionality and factorial composition of the AQ scale was investigated by exploratory and confirmatory factor analyses. Since spurious associations and increased factorial complexity can occur when the analysis fails to consider the inherently categorical nature of the item level data, this study, in contrast to most previous studies, estimated the correlation matrices subjected to factor analysis using the polychoric correlations. The resulting factors were validated via sociodemographic characteristics and psychopathological scales obtained from the respondents.*

*The results showed that based on the distribution of factor loadings and factor correlations, in the entire nationally representative sample of 1200 adult subjects, from the original factor structure three of the four factors (Physical and Verbal Aggression and Hostility) showed a good replication whereas the fourth factor (Anger) replicated moderately well. Replication further improved when the sample was restricted in age, i.e. the analysis focused on a sample representing the younger age group, comparable to that used in the original Buss–Perry study. Similar to the Buss–Perry study, and other investigations of the AQ scale, younger age and male gender were robustly related to physical aggression. In addition, level of verbal aggression was different between the two genders (with higher severity in males) whereas hostility and anger were essentially the same in both genders.*

*In conclusion, the current study based on a representantive sample of adult population lends support to the use of the AQ scale in the general population. The authors suggest to exclude from the AQ the two inverse items because of the low reliability of these items with regard to their hypothesized constructs. Copyright © 2007 John Wiley & Sons, Ltd.*

**Key words:** Buss and Perry's Aggression Questionnaire, generalizability, exploratory and confirmatory factor analyses, representative sample

**Introduction**

Despite of the widespread application of a self-report instrument, the 29-item Buss–Perry Aggression Questionnaire (AQ) (Buss and Perry, 1992), several critical issues with regard to the potential application of the scale remain. These issues include the generalizability of the results based on (1) the narrowly selected samples that have been subjected to testing and (2) the application of methods that fail to take into consideration the psychometric properties of the scale.

First, the validation of the scale was examined only in restricted samples which were not representative of an adult population (Harris, 1995; Harris and Knight-Bohnhoff, 1996; Lovaš and Trenková, 1996; Meesters et al., 1996; Bernstein and Gesn, 1997; Ando et al., 1999; Felsten and Hill, 1999; Bryant and Smith, 2001; Nakano, 2001; Ramirez et al., 2001; García-Leon et al., 2002; Morren and Meesters, 2002; Fossati et al., 2003; Gallardo et al., 2004; Collani and Werner, 2005; Palmer and Thakordas, 2005; Vigil-Colet et al., 2005). This creates a problem because individuals, who are at risk of displaying aggressive and hostile behaviours are not restricted to those samples that have been subjected to psychometric testing regarding the application of the scale (e.g. the prison population, which typically displays a high level of aggression (Van Praag et al., 1990).

In particular, based on our systematic review of literature, the samples studied in previous investigations were characterized by the preponderance of younger age group and female subjects as well as a high education level of respondents. In the majority of the studies, the examined special populations are not suitable to yield universally generalizable results. For example, based on published findings, females are less aggressive than men (Volavka, 2002), the risk period of violence develops by the end of the young adulthood (Hollander and Stein, 1995), and individuals with higher education are less violent than those with lower education (Volavka, 2002). We summarized the characteristics of the samples used in the previous studies in Table 1.

The aforementinoned problem – i.e. potential lack of generalizability – has recently been recognized in the literature. For example, Collani and Werner (2005) pointed out the problem with the generalization of the AQ results to a broader population. These authors attempted to overcome this problem by recruiting subjects via the internet, by 'approaching individuals from a general adult population'. However, as the authors pointed out their sample was 'severely biased' as well, with very few participants from groups with lower education and from lower social class. Similar to Collani and Werner, Vigil-Colet et al. (2005) stressed the importance of studying the factorial structure of the AQ scale in more heterogenous populations instead of examining secondary school and university students.

Second, as has been pointed out by several authors, the AQ scale is based on a five-item categorical rating, and in general has a skewed distribution. However, the overwhelming majority of the analyses used normal theory approaches (e.g. use of Pearson correlations) to describe associations among the individual items of the scale. As a result, spurious associations and increased factorial complexity can occur when the analysis fails to consider the inherently categorical nature of the item level data (e.g. the number of factors as determined from the analyses is greater than the true number of factors that underlie the scale). As pointed out by Bernstein and Gesn (1997, p. 564), such an effect (i.e. spurious inflation of dimensionality) is attributable to the fact that 'items with similar distributions tend to correlate more highly with one another than items with dissimilar distributions. As a result, items tend to form factors on the basis of the similarity of their item distributions as well as their content. The effect is basically a form of range restriction'.

In addition, in previous exploratory and confirmatory factor analyses the resulting factor structures accounted for only relatively small amounts of variance. To improve this situation, previous investigators used two approaches: (a) restriction of the items of the scale to a smaller subset (e.g. the reduced 12-item version of the scale, introduced by Bryant and Smith (2001) and (b) omission of certain items that have low indicator reliability with regard to their factors (e.g. to reversely scaled items, one in the Physical Aggression, and one in the Verbal Aggression factors).

Based on the problems that we identified in the previous studies, the principal objective of this study was to examine the factorial structure and the psychometric properties of the AQ scale in a nationally representative sample of the Hungarian adult population. In addition, in order to overcome the potential problems associated with the analyses of categorical data with normal theory approaches, we estimated the correlation matrices subjected to factor analysis using the polychoric correlations. Furthermore, we examined whether reduction of the original 29-item scale to a



DOI: 10.1002/mpr

**Table 1.** Number of observations and main characteristics of samples used in previous studies

| Author, Date | Target population | Age | Number of observations | Number (ratio) of females | Version of AQ |
|---|---|---|---|---|---|
| Buss and Perry, 1992 | College students (USA) | Average age: 18–20 years | *N* = 1253, three consecutive samples of 406, 448, 399, | 641 (0.52) | 29-item version |
| Harris, 1995 | University students (USA) | Not provided | *N* = 306 | 151 (0.49) | 29-item version |
| Meesters et al., 1996 | University students (The Netherlands) | Average age 21.4 | *N* = 762 | 518 (0.68) | 29-item version |
| Harris and Knight-Bohnhoff, 1996 | College students (USA)<br>Military workers (USA) | Not provided | Total *N* = 194<br>*N* = 115 (students)<br>*N* = 79 (military workers) | 73 (0.38) | 29-item version |
| Bernstein and Gesn, 1997 | University students (USA) | From the late teens to the early forties | *N* = 321 | 208 (0.65) | 29-item version |
| Felsten and Hill, 1999 | Undergraduate students (USA) | 23.6 | *N* = 36 | 21 (0.58) | 29-item version |
| Bryant and Smith, 2001 | Undergraduate students | | | | 29-item |
| | USA | 18.94 | *N* = 307 | 173 (0.56) | version |
| | UK | 25.13 | *N* = 200 | 100 (0.50) | 12-item |
| | Canada | Not provided | *N* = 306 | 151 (0.49) | version |
| | USA | 18.35 | *N* = 171 | 123 (0.72) | |
| | USA | 18.64 | *N* = 170 | 124 (0.73) | |
| Ramirez et al., 2001 | Undergraduate students | | | | 29-item |
| | Japan | 20.50 | *N* = 200 | 100 (0.50) | version |
| | Spain | 19.00 | *N* = 200 | 136 (0.68) | |
| Nakano, 2001 | Undergraduate students (Japan) | 18–24 | *N* = 425 | 256 (0.60) | 29-item version |
| Fossatti et al., 2003 | University students (Italy) | 23.41 | *N* = 392 | 237 (0.60) | 29-item version |
| Gallardo et al., 2004 | Subjects (Italy) | 33.99 | *N* = 1268 | Not provided | 12-item version |
| Collani and Werner, 2005 | Mixed population | 28.6 | *N* = 417 | 288 (0.69) | 29-item |
| | Student population | Not provided | *N* = 141 | 116 (0.82) | version |
| | Mixed population (Germany) | 34.7 | *N* = 401 | 217 (0.54) | |

12-item version, as proposed in the literature, would result in a better fit of the postulated factors to the empirical data.

## Methods

*Sample*

The sample for the study comprised 1200 persons representative of the population over 18 years of age selected by the two-step, group-stratified sampling method. In the first step, a representative sample of geographic areas (neighbourhouds) were selected on the basis of the size. In the second step the respondents were randomly sampled using a schedule of probability sampling. In the current sample, the refusal rate was approximately 20.3%. The rate of refusal was not a function of potentially important demographic variables such a gender or age. In females and males, refusal



DOI: 10.1002/mpr

rates were 19.5% and 21.2%, respectively; for three age groups including low (18–35 years), medium (36–55 years), and higher ages (above 55 years) refusal rates were 21.5%, 19.6% and 19.9%, respectively.

Overall, 638 females and 562 males consented to participate in the survey that we report here. The average age was 46.6 years (standard deviation (SD) = 18.1). Eighteen per cent of the respondents lived in Budapest, 35% in a rural environment, and 47% in cities. Altogether, 11.7% had a high level of education, whereas 57.3% had a lower level of education (ending the secondary school). The selected sample represented 0.015% of the entire adult population in Hungary.

*Data collection*

Data collection was performed in two weekends in April 2005, as part of a research project entitled 'Public Opinion of the Relationships Between Alcohol Problems and Criminality' (Pygmalion Project, 2004–2007). Data for the analyses were collected at the home of the participants through a face-to-face interview and a self-report questionnaire. The administration of tests to respondents took an average time of 55 minutes. The interviews were carried out by previously trained interviewers, who had prior experience in social and clinical research.

*Measures*

The AQ comprises 29 items of a five-point Likert format from one ('extremely uncharacteristic of me') to five ('extremely characteristic of me'). The four widely used subscales of the questionnaire, established on the basis of factor analyses, were Physical Aggression (PA, nine items), Verbal Aggression (VA, five items), Anger (AN, seven items) and Hostility (HS, eight items). A Hungarian version of the AQ was translated from English by two bilingual individuals, and then independently translated back to English items by another bilingual individual. Any discrepancy that occurred was resolved on the basis of a consensus decision.

The Rosenberg Self-esteem (RSE) scale is a unidimensional measure tool of global self-esteem. Rosenberg scored his 10-question scale that was presented with four response choices, as a six-item Guttman scale, as ranging from strongly agree to strongly disagree (Rosenberg, 1965). For the purpose of this investigation, the scale was represented by two subscales, the sum of positively and negatively scaled items (i.e. high and low self esteem), respectively.

The shortened version of Beck Depression Inventory (S-BDI), originally designed for the measurement of the severity of depression in patients with psychiatric diagnoses, includes 13 items, which are rated on a four-point severity scale from 0 to 3 (Beck and Beck, 1972).

*Analyses*

As mentioned earlier, the principal purpose of this study was to examine the factorial structure and the psychometric properties of the AQ scale in a nationally representative sample of the Hungarian adult population and to investigate whether the factor structure described originally by Buss and Perry generalize to this population. For this investigation, generalizability was viewed as factorial invariance, i.e. constancy in the structure of the underlying factors across samples (current sample versus original sample from the Buss and Perry–Perry investigation). The concept of factorial invariance is built on Thurstone's notion of simple structure, which posits that the pattern of salient (non-zero) and non-salient (zero or near-zero) loadings defines the structure of a psychometric construct (Thurstone, 1947). With regard to factorial invariance, the notion of simple structure stipulates configurational invariance; items associated with the same construct are expected to exhibit the same configuration of salient and non-salient factor loadings across groups (or conditions).

The factor analyses were conducted in two steps. First, the empirical data from the current sample were subjected to exploratory factor analysis (EFA) to examine whether model modifications were needed in terms of the number of the factors and item composition of the underlying factors derived from the sample in the Buss–Perry investigation. Second, confirmatory factor analyses (CFA) were performed to test the configurational invariance of the hypothesized factor structure, i.e. to examine whether the items have the same relationship to the same underlying factor as posited on the basis of the earlier analyses in from the Buss–Perry sample. The analyses in this study were based on the polychoric (instead of the Pearson's) correlations in order to account for the categorical nature of the data.

Step 1 – exploratory factor analyses (EFA)

The purpose of the EFA was to investigate whether the theoretically-postulated factor structure derived from Buss and Perry's sample represents an adequate



DOI: 10.1002/mpr

representation of the pattern of observed associations among scale items in our study. More specifically, in these analyses we investigated whether model improvements were necessary regarding the number of factors, and the factor structure of the individual factors based on the pattern of salient and non-salient loadings. Similar to Buss and Perry's investigation, the principal component method was used for factor extraction, and the OBLIMIN rotation was applied for deriving a simple structure to help the interpretation. The OBLIMIN rotation (Loehlin, 1987) is an oblique rotation technique which permits correlation between factors. Since there are conceptual as well as clinical reasons to presume a correlation between the factors of the AQ scale, this technique provides a more realistic representation of the data than the orthogonal solution which assumes independence. For the examination of the dimensionality in an EFA, we used the Kaiser–Guttman eigenvalue >1 criterion (Kaiser, 1960) as well as Cattell's Scree plot (Cattell, 1978). Items were allocated to factors according to their highest loading; similar to Buss and Perry's investigation, the threshold loading of 0.35 was chosen to indicate saliency. Degree of congruence ('match') between pairs of factors derived from the Buss and Perry sample and the current sample, respectively, was expressed by the coefficient of congruence (Kaiser, 1960), which expresses similarity based on both the pattern and the magnitude of loadings, ranging from ±1.0 (perfect/inverse agreement) to 0.0 (no agreement). In addition, correlation between pairs of factor scores computed on the basis of the Buss and Perry factors (specified a priori based on Buss and Perry's results) and those derived empirically from the current sample were also used to investigate similarity between the two factor solutions.

Step 2 – confirmatory factor analyses (CFA)

The relationship between the observed variables and the hypothesized underlying constructs was investigated by the CFA model. Based on the findings from the Buss and Perry sample, the CFA model used in this investigation posited an a priori definition of the factor structure (measurement model). The four factors were considered as interrelated constructs in the structural part of the CFA models; correlation was therefore allowed between any two of the factors. In the CFA, estimates of loadings of the individual items on the AQ scale were obtained for their hypothesized factors. Values of *t*-statistics were applied to test whether the individual items were significantly related to their specific factors.

The root mean square error of approximation (RMSEA) and the goodness-of-fit index (GFI) were used to assess model fit for the entire CFA model. This measure indexes the fit of the model to the covariance matrix. It represents the square root of the average amount that the sample covariances differ from their estimates derived from the stipulated factor structure. As a guide, RMSEA values below 0.1 are regarded as an adequate fit, whereas values of <0.05 are considered as a close fit. With regard to GFI, values above 0.90 are viewed as a reflection of an adequate model fit.

Reliability, construct validity

Scale (factorial) reliability was investigated through the internal consistency reliability. The Cronbach alpha coefficient (Cronbach, 1951) was used to determine the internal consistency for each of the four factors of the AQ scale. External (criterion-related) validity of the four factors was investigated via relationships with demographic variables (gender, age) that have been shown to be related in previous studies to individual factors of the AQ scale. In addition, the degree to which the four factors of the AQ scale overlapped with ratings of psychopathology was also investigated. Specifically, association between each of the Buss–Perry factors and the individual subscales of the Beck depression scale and the Rosenberg scale was investigated using bivariate correlations.

Ancillary analyses

The objective of these analyses was three-fold. First, since the original Buss and Perry study focused on a sample of younger individuals, we wanted to explore whether the restriction of the sample to a younger age group increases the replicability. Second, based on Bryant–Smith's study indicating that the original 29-version of the AQ scale can be reduced to a substantially smaller set of items ($N = 12$), we examined whether restriction of the items to this subset would allow for a reliable the replication of four-factor structure of AQ scale. Third, since the AQ scale comprises two inversely scaled items which may have low reliablity and internal consistency with the rest of the items, we examined whether the omission of these items from the AQ increases the model fit in the CFA analysis.

We used SAS (the Statistical Analysis System, version 9.1) to perform the factor analyses (EFA, CFA



DOI: 10.1002/mpr

implemented with FACTOR and CALIS procedures, respectively) and to carry-out the validation [implemented with the General Linear Model (GLM) procedure].

## Results

*Basic Demographic Characteristics and Descriptive Statistics*

The mean age of the target population of this study was 46.6 (SD = 0.52) and the gender distribution indicated a slightly higher ratio of females in the sample (55.2%), reflecting the characteristics of the general population.

As shown in Table 2, the mean item score across the entire range of items was approximately two, which yielded a total score of 57.2 (SD = 14.9). The mean score for the individual items ranged between 1.37 (SD = 0.68) for item 1 (PA) and 3.58 (SD = 0.99) for item 3 on the AN subscale.

*Exploratory factor analysis (EFA)*

The EFA analysis of the full sample of subjects yielded four factors based both on the Kaiser's eigenvalue greater than one criterion and on the scree-plot. Altogether, the four factors explained 38.4%, 8.5%, 7.0%, 4.7% of the total variance of the 29 items of the scale, respectively. The four factors, taken together, explained 58.5% of the variance.

The results of exploratory factor analysis are displayed in Table 3. The first factor that we identified had high loadings (>0.35) on items of PA, indicating a good replication of the original factor reported by Buss and Perry. The second factor that we identified as HS obtained high item loadings on the HS subscale with the exception of item 1 ('I am sometimes eaten up with jealousy'). The third factor showed high loadings on items of VA. The fourth factor (AN) indicated only a moderate replication, with a majority of items loading on AN or HS. The correlation coefficients between the factor scores of the a priori defined Buss–Perry factors and empirical factors were 0.86, 0.92, 0.86 and 0.63 for the PA, HS, VA and AN factors, respectively. The coefficients of congruence between the factor loadings of the original Buss and Perry factors, and the loadings of the empirically determined factors from our study were 0.77, 0.79, 0.58 and 0.39 for the PA, HS, VA and AN factors, respectively. We note that in the current sample the order of the four factors based on the explained variance was different (1 = PA, 2 = HS, 3 = VA, 4 = AN) compared to the order of factors based on the original Buss and Perry's study.[1]

*Confirmatory factor analysis (CFA)*

The fit indices for confirmatory factor model indicated a moderate fit with a GFI value of approximately 0.82 and AGFI of 0.79. The root mean square residual was 0.07 which, similar to a GFI and AGFI (Adjusted Goodness of Fit Index) suggested a moderate fit.

The loadings that we derived from the CFA model based on the a priori defined factor structure are displayed in Table 4. Columns one and two display the estimated values for the loadings and the corresponding standard error respectively. As Table 4 shows, the posited factor structure yielded a statistical significant value (loading) for each of the items and for each of the postulated factors. However consistent with the results from the EFA analyses, the loadings were typically higher for PA and HS as compared to the two remaining subscales (VA, AN). We note that for the two inversely scaled items ('I am an even-tempered person'; 'I can think of no good reason for ever hitting a person') the loading values, were low, despite the fact that they acquired statistical significance.

*Internal consistency reliability*

Investigation of the internal consistency reliability was based on the Cronbach coefficient alpha for each of the four individual factors. The results are shown in Table 5. The alpha values depicted in Table 5 indicate high internal consistency for two of the factors including PA and HS, whereas for VA and AN the reliabilities were moderate.

*External validation*

The comparison of the two genders based on the four empirically derived factors indicated a statistically significant difference for PA and VA, whereas no difference was found for HS and AN. Results of the gender comparisons are displayed in Table 6. The effect sizes displayed in Table 6 indicate strong association between PA and male gender (Cohen's $d = 0.6$), whereas for VA the association was weak (Cohen's $d = 0.22$). With regard to age, a statistically significant but modest negative association was found for three of the four factors including PA, HS and VA).

Factor 1 on the Rosenberg scale showed a significant correlation with AN on the AQ scale



DOI: 10.1002/mpr

**Table 2.** Basic descriptive statistics for each item on the Buss–Perry AQ

| Subscales and items of the AQ[1] | Mean ($N$ = 1200)[2] | SD | Q1–Q3[3] |
|---|---|---|---|
| *Physical Aggression (PA)* | | | |
| 1. Once in a while I can't control the urge to strike another person. | 1.37 | 0.68 | 1.00–2.00 |
| 2. Given enough provocation, I may hit another person. | 1.52 | 0.79 | 1.00–2.00 |
| 3. If somebody hits me, I hit back. | 2.21 | 1.21 | 1.00–3.00 |
| 4. I get into fights a little more than the average person. | 1.35 | 0.66 | 1.00–2.00 |
| 5. If I have to resort to violence to protect my rights, I will. | 1.93 | 1.03 | 1.00–2.00 |
| 6. There are people who pushed me so far that we came to blows. | 1.44 | 0.74 | 1.00–2.00 |
| 7. I can think of no good reason for ever hitting a person.[4] | 3.01 | 1.49 | 2.00–4.00 |
| 8. I have threatened people I know. | 1.53 | 0.78 | 1.00–2.00 |
| 9. I have become so mad that I have broken things. | 1.50 | 0.79 | 1.00–2.00 |
| *Verbal Aggression (VA)* | | | |
| 1. I tell my friends openly when I disagree with them. | 3.41 | 1.12 | 3.00–4.00 |
| 2. I often find myself disagreeing with people. | 2.53 | 0.94 | 2.00–3.00 |
| 3. When people annoy me, I may tell them what I think of them. | 3.01 | 1.14 | 2.00–4.00 |
| 4. I can't help getting into arguments when people disagree with me. | 2.00 | 1.01 | 1.00–3.00 |
| 5. My friends say that I'm somewhat argumentative. | 1.56 | 0.74 | 1.00–2.00 |
| *Anger (AN)* | 2.36 | 1.16 | 1.00–3.00 |
| 1. I flare up quickly but get over it quickly. | 2.80 | 1.07 | 2.00–4.00 |
| 2. When frustrated, I let my irritation show. | 1.93 | 0.99 | 1.00–2.00 |
| 3. I sometimes feel like a powder keg ready to explode | 3.58 | 0.99 | 3.00–4.00 |
| 4. I am an even-tempered person.[4] | 1.81 | 0.95 | 1.00–2.00 |
| 5. Some of my friends think I'm a hothead. | 1.78 | 0.87 | 1.00–2.00 |
| 6. Sometimes I fly off the handle for no good reason. | 1.73 | 0.78 | 1.00–2.00 |
| 7. I have trouble controlling my temper. | 2.36 | 1.16 | 1.00–3.00 |
| *Hostility (HS)* | | | |
| 1. I am sometimes eaten up with jealousy. | 1.73 | 0.84 | 1.00–2.00 |
| 2. At times I feel I have gotten a raw deal out of life. | 2.93 | 1.17 | 2.00–4.00 |
| 3. Other people always seem to get the breaks. | 1.67 | 0.78 | 1.00–2.00 |
| 4. I wonder why sometimes I feel so bitter about things. | 1.96 | 0.94 | 1.00–2.00 |
| 5. I know that 'friends' talk about me behind my back. | 1.84 | 0.92 | 1.00–2.00 |
| 6. I am suspicious of overly friendly strangers. | 2.58 | 1.08 | 2.00–3.00 |
| 7. I sometimes feel that people are laughing at me behind my back. | 1.78 | 0.86 | 1.00–2.00 |
| 8. When people are especially nice, I wonder what they want. | 2.16 | 1.02 | 1.00–3.00 |
| AQ total[5] | 57.19 | 14.89 | 47.00–66.00 |

[1] Buss–Perry Aggression Questionnaire.
[2] Sample size for individual items may vary due to missing data.
[3] Q1–Q3 = Interquartile Range.
[4] Items scored positively with increasing value indicating better functioning.
[5] For total score computations scoring was reversed for two positively scored items.

(Pearson $r = -0.44$, $p < 0.0001$). Factor 2 (negatively scaled items) had significant association with HS (Pearson $r = 0.45$, $p < 0.0001$), VA (Pearson $r = 0.25$, $p < 0.0001$) and AN (Pearson $r = 0.47$, $p < 0.0001$). The first factor on the Beck scale (psychosomatic subscale) showed a significant correlation with AN on the AQ scale (Pearson $r = 0.28$, $p < 0.0001$). Factor 2 (dysthymia) had significant association with HS (Pearson $r = 0.32$, $p < 0.0001$) and AN (Pearson $r = 0.42$, $p < 0.0001$).



DOI: 10.1002/mpr

**Table 3.** Exploratory factor analysis (EFA). Factor loadings

| Exploratory factor[1] | Items of the AQ[2] | Factor loading ($N = 1200$)[3] | | | |
|---|---|---|---|---|---|
| | | Factor 1 (PA) | Factor 2 (HS) | Factor 3 (VA) | Factor 4 (AN) |
| Physical Aggression (PA) | 1. Once in a while I can't control the urge to strike another person. | 0.61* | 0.24 | 0.04 | 0.29 |
| | 2. Given enough provocation, I may hit another person. | 0.68* | 0.18 | 0.21 | 0.10 |
| | 3. If somebody hits me, I hit back. | 0.60* | 0.08 | 0.38 | −0.06 |
| | 4. I get into fights a little more than the average person. | 0.62* | 0.10 | 0.07 | 0.39* |
| | 5. If I have to resort to violence to protect my rights, I will. | 0.55* | 0.39 | 0.22 | −0.19 |
| | 6. There are people who pushed me so far that we came to blows. | 0.68* | 0.13 | 0.16 | 0.26 |
| | 7. I can think of no good reason for ever hitting a person.[4] | −0.35* | −0.14 | 0.28 | −0.10 |
| | 8. I have threatened people I know. | 0.47* | 0.33 | 0.05 | 0.28 |
| | 9. I have become so mad that I have broken things. | 0.47* | 0.24 | 0.12 | 0.31 |
| Verbal Aggression (VA) | 1. I tell my friends openly when I disagree with them. | −0.03 | 0.25 | 0.62* | −0.48* |
| | 2. I often find myself disagreeing with people. | −0.27 | 0.31 | 0.42* | 0.20 |
| | 3. When people annoy me, I may tell them what I think of them. | 0.07 | 0.13 | 0.75* | −0.23 |
| | 4. I can't help getting into arguments when people disagree with me. | 0.27 | 0.27 | 0.45* | 0.06 |
| | 5. My friends say that I'm somewhat argumentative. | 0.30 | 0.14 | 0.33 | 0.44* |
| Anger (AN) | 1. I flare up quickly but get over it quickly. | 0.13 | 0.20 | 0.56* | 0.14 |
| | 2. When frustrated, I let my irritation show. | −0.08 | 0.04 | 0.74* | 0.14 |
| | 3. I sometimes feel like a powder keg ready to explode. | 0.10 | 0.24 | 0.40* | 0.38* |
| | 4. I am an even-tempered person.[4] | −0.03 | −0.05 | 0.27 | −0.69* |
| | 5. Some of my friends think I'm a hothead. | 0.43* | 0.22 | 0.28 | 0.05 |
| | 6. Sometimes I fly off the handle for no good reason. | 0.11 | 0.02 | 0.42* | 0.53* |
| | 7. I have trouble controlling my temper. | 0.22 | 0.13 | 0.34 | 0.44* |
| Hostility (HS) | 1. I am sometimes eaten up with jealousy. | 0.00 | 0.11 | 0.12 | 0.45* |
| | 2. At times I feel I have gotten a raw deal out of life. | −0.48* | 0.39* | 0.30 | 0.31 |
| | 3. Other people always seem to get the breaks. | 0.13 | 0.34 | 0.18 | 0.48* |
| | 4. I wonder why sometimes I feel so bitter about things. | −0.14 | 0.61* | 0.03 | 0.25 |
| | 5. I know that 'friends' talk about me behind my back. | −0.02 | 0.46* | 0.05 | 0.41* |
| | 6. I am suspicious of overly friendly strangers. | −0.06 | 0.83* | −0.06 | −0.34 |
| | 7. I sometimes feel that people are laughing at me behind my back. | −0.14 | 0.54* | 0.00 | 0.48* |
| | 8. When people are especially nice, I wonder what they want. | 0.08 | 0.88* | −0.06 | −0.24 |

[1] For the purpose of EFA items were organized based on Buss and Perry (1992). EFA was based on the Principal Component Solution with Oblimin Rotation.
[2] Buss–Perry Aggression Questionnaire.
[3] Sample size for individual items may vary due to missing data.
[4] Items scored were positively with increasing value indicating better functioning.
* Loading ≥ 0.35.



DOI: 10.1002/mpr

**Table 4.** Confirmatory factor analysis (CFA). Estimates of factor loadings

| Confirmatory factor[1] | Items of the AQ[2] | Loading ($N = 1200$)[3] | Standard error | *t*-Statistic[4] |
|---|---|---|---|---|
| Physical Aggression (PA) | 1. Once in a while I can't control the urge to strike another person. | 0.718 | 0.027 | 26.65 |
| | 2. Given enough provocation, I may hit another person. | 0.779 | 0.026 | 29.90 |
| | 3. If somebody hits me, I hit back. | 0.616 | 0.028 | 21.83 |
| | 4. I get into fights a little more than the average person. | 0.724 | 0.027 | 26.97 |
| | 5. If I have to resort to violence to protect my rights, I will. | 0.590 | 0.029 | 20.68 |
| | 6. There are people who pushed me so far that we came to blows. | 0.786 | 0.026 | 30.33 |
| | 7. I can think of no good reason for ever hitting a person.[5] | −0.193 | 0.031 | −6.16 |
| | 8. I have threatened people I know. | 0.680 | 0.027 | 24.80 |
| | 9. I have become so mad that I have broken things. | 0.644 | 0.028 | 23.11 |
| Verbal Aggression (VA) | 1. I tell my friends openly when I disagree with them. | 0.286 | 0.032 | 9.08 |
| | 2. I often find myself disagreeing with people. | 0.445 | 0.031 | 14.57 |
| | 3. When people annoy me, I may tell them what I think of them. | 0.486 | 0.030 | 16.07 |
| | 4. I can't help getting into arguments when people disagree with me. | 0.643 | 0.029 | 22.24 |
| | 5. My friends say that I'm somewhat argumentative. | 0.068 | 0.029 | 23.68 |
| Anger (AN) | 1. I flare up quickly but get over it quickly. | 0.684 | 0.028 | 24.66 |
| | 2. When frustrated, I let my irritation show. | 0.580 | 0.029 | 20.02 |
| | 3. I sometimes feel like a powder keg ready to explode. | 0.699 | 0.028 | 25.37 |
| | 4. I am an even-tempered person.[5] | −0.233 | 0.032 | −7.39 |
| | 5. Some of my friends think I'm a hothead. | 0.544 | 0.029 | 18.52 |
| | 6. Sometimes I fly off the handle for no good reason. | 0.643 | 0.028 | 22.75 |
| | 7. I have trouble controlling my temper. | 0.652 | 0.028 | 23.14 |
| Hostility (HS) | 1. I am sometimes eaten up with jealousy. | 0.373 | 0.032 | 11.81 |
| | 2. At times I feel I have gotten a raw deal out of life. | 0.470 | 0.031 | 15.21 |
| | 3. Other people always seem to get the breaks. | 0.680 | 0.029 | 23.69 |
| | 4. I wonder why sometimes I feel so bitter about things. | 0.561 | 0.030 | 18.68 |
| | 5. I know that 'friends' talk about me behind my back. | 0.640 | 0.029 | 21.96 |
| | 6. I am suspicious of overly friendly strangers. | 0.320 | 0.032 | 10.03 |
| | 7. I sometimes feel that people are laughing at me behind my back. | 0.665 | 0.029 | 23.04 |
| | 8. When people are especially nice, I wonder what they want. | 0.476 | 0.031 | 15.43 |

[1] For the purpose of CFA four factors were posited based on Buss and Perry (1992).
[2] Buss–Perry Aggression Questionnaire.
[3] Sample size for individual items may vary due to missing data.
[4] $p < 0.05$ (nominal value) for all entries in the column.
[5] Items scored were positively with increasing value indicating better functioning.



DOI: 10.1002/mpr

**Table 5.** Internal consistency reliability (Cronbach alpha) and item composition of each factor on the Buss–Perry AQ

| Factor | Items of the AQ[1] | Cronbach alpha ($N = 1200$)[2] | |
|---|---|---|---|
| | | Alpha without specific item | Overall alpha for factor for all items in factor |
| Physical Aggression (PA) | 1. Once in a while I can't control the urge to strike another person. | 0.788 | 0.82 |
| | 2. Given enough provocation, I may hit another person. | 0.779 | |
| | 3. If somebody hits me, I hit back. | 0.799 | |
| | 4. I get into fights a little more than the average person. | 0.788 | |
| | 5. If I have to resort to violence to protect my rights, I will. | 0.802 | |
| | 6. There are people who pushed me so far that we came to blows. | 0.779 | |
| | 7. I can think of no good reason for ever hitting a person.[3] | 0.878 | |
| | 8. I have threatened people I know. | 0.794 | |
| | 9. I have become so mad that I have broken things. | 0.796 | |
| Verbal Aggression (VA) | 1. I tell my friends openly when I disagree with them. | 0.639 | 0.68 |
| | 2. I often find myself disagreeing with people. | 0.648 | |
| | 3. When people annoy me, I may tell them what I think of them. | 0.590 | |
| | 4. I can't help getting into arguments when people disagree with me. | 0.588 | |
| | 5. My friends say that I'm somewhat argumentative. | 0.658 | |
| Anger (AN) | 1. I flare up quickly but get over it quickly. | 0.603 | 0.70 |
| | 2. When frustrated, I let my irritation show. | 0.627 | |
| | 3. I sometimes feel like a powder keg ready to explode. | 0.622 | |
| | 4. I am an even-tempered person.[3] | 0.600 | |
| | 5. Some of my friends think I'm a hothead. | 0.659 | |
| | 6. Sometimes I fly off the handle for no good reason. | 0.640 | |
| | 7. I have trouble controlling my temper. | 0.638 | |
| Hostility (HS) | 1. I am sometimes eaten up with jealousy. | 0.753 | 0.75 |
| | 2. At times I feel I have gotten a raw deal out of life. | 0.732 | |
| | 3. Other people always seem to get the breaks. | 0.717 | |
| | 4. I wonder why sometimes I feel so bitter about things. | 0.718 | |
| | 5. I know that 'friends' talk about me behind my back. | 0.710 | |
| | 6. I am suspicious of overly friendly strangers. | 0.748 | |
| | 7. I sometimes feel that people are laughing at me behind my back. | 0.699 | |
| | 8. When people are especially nice, I wonder what they want. | 0.723 | |

[1] Buss–Perry Aggression Questionnaire.
[2] Sample size for individual items may vary due to missing data.
[3] Items scored were positively with increasing value indicating better functioning.

*Ancilliary analyses*

Subgroup analysis

The objective of this analysis was to explore whether the restriction of the sample to a younger age group which was the focus of the original Buss–Perry study enhances the replicability of our findings. To accomplish this, we re-ran the EFA and CFA analyses with a sample of subjects of 30 years or younger. The order of the four factors based on the explained variance was comparable (1 = PA, 2 = VA, 3 = HS, 4 = AN) to that of the original study (Buss and Perry, 1992). The correlation coefficients between the factor scores of the a priori defined Buss and Perry factors and the empirical



DOI: 10.1002/mpr

**Table 6.** Criterion-related validity. Difference between genders on four of the empirically derived factors of Buss–Perry AQ

| Factors from EFA | Gender Factor scores[2], mean (SD) | | | | Effect size (Cohen *d*) | Test-statistic[1] | |
|---|---|---|---|---|---|---|---|
| | Male ($n$ = 538) | | Female ($n$ = 662) | | | *F* | *p* |
| BP Factor 1 (PA) | 0.290 | (0.99) | −0.244 | (0.80) | 0.60 | 99.29 | <0.0001* |
| BP Factor 2 (HS) | 0.037 | (0.93) | −0.039 | (0.93) | 0.08 | 1.83 | 0.18 (ns) |
| BP Factor 3 (VA) | 0.123 | (0.97) | −0.086 | (0.95) | 0.22 | 13.00 | 0.0003* |
| BP Factor 4 (AN) | −0.004 | (1.01) | −0.017 | (0.85) | 0.03 | 0.005 | 0.82 (ns) |

[1]GLM analysis.
[2]Raw mean (SD) factor scores for empirically derived factors on the Buss–Perry AQ.
*$p < 0.05$.
Note: ns, not significant.

factors were 0.90, 0.94, 0.87 and 0.77 for the PA, HS, VA and AN factors, respectively. Thus, these results indicated an increase in replicability when the sample was similar to that used in the Buss and Perry's study.

12-item version

Bryant and Smith's (2001) study indicated that the original four-factor structure of AQ scale can be replicated reliably with a substantially reduced set of items ($N$ = 12). To examine this issue, based on the Bryant–Smith's study we included 12 items in the analysis, and re-ran the CFA analysis. Similar to Brian and Smith's study, the results revealed substantial improvement in the fit of the CFA factor model compared to the analysis of the original 29-item scale (GFI = 0.94, AGFI = 0.91, RMSE = 0.05). However, additional inspection of the results indicated that this increase in model fit, at least in the current study, was attributable to a singular covariance structure (Heywood Case) due to the close association between the 12 items selected for the analysis.

Omission of inversely scaled items from the questionnaire

The goal of this analysis was to examine whether the omission of the two inversely scaled items from the questionnaire increases the model fit as determined from the CFA analysis. In order to do this, we re-ran the CFA analysis with 29 items using the same a priori defined factors as in the analysis of the full set of items. As expected, the results showed a moderate improvement in model fit (GFI = 0.84, AGFI = 0.81, RMSE = 0.074), indicating, that these items have low reliability with the rest of the items.

## Discussion

Our study investigated the factorial structure and psychometric properties of the Buss and Perry AQ in a nationally representative sample of the Hungarian adult population (over 18 years of age). In contrast to previous studies, where younger age groups, females and higher education were represented at a disproportionately high rate compared to a general population (Buss and Perry, 1992; Harris, 1995; Harris and Knight-Bohnhoff, 1996; Meesters et al., 1996; Bernstein and Gesn, 1997; Felsten and Hill, 1999; Bryant and Smith, 2001; Nakano, 2001; Ramirez et al., 2001; García-León et al., 2002; Morren and Meesters, 2002; Fossati et al., 2003; Collani and Werner, 2005), this sample was representative for age, gender, education and population density in the adult population.

Results of our analyses indicated that, similar to Buss and Perry's findings, the four-factor structure provided an adequate representation of the data. Furthermore, in terms of variance explained by the factors, our results are comparable with earlier factor analyses of the Buss–Perry scale. For example, using a large sample of normal subjects as Bernstein and Gesn (1996) reported that four factors, taken together, explained approximately 47.3% of the variation of the 29-item version of the Buss–Perry's scale. Additionally, based on the pairs of correlations between the factor sores of empirically determined factors from our sample and those defined a priori based on Buss and Perry's results, three of the factors (PA, HS, VA) showed a good, whereas the fourth factor (AN) showed a moderate replication. Results from the analyses of the coefficients of congruence provided results consistent with those derived from the correlations.



DOI: 10.1002/mpr

Similar to the result from the EFA, the CFA showed the four-factor structure, defined a priori based on Buss and Perry's data, was therefore replicable in a representative sample. In addition, the inspection of the factor loadings from the CFA, showed the best replicability for the PA, HS and VA factors, whereas the replicability of the AN factor was only modest. Consistent with previous studies, the examination of the fit indices (GFI, AGFI, RMSE) indicated the four-factor structure can provide only moderate fit to the data.

In order to examine whether the correspondence between our results and those published in the literature can be further increased, we conducted more detailed analyses with regard to three issues raised previously in the literature.

First, we investigated whether by restricting our study population to a sample with a younger age, the factor structure derived from our study will show a higher correspondence with that from the Buss and Perry's study that was based on a sample of subjects with younger age. Results from both the EFA and CFA indicated a better overall replication of the data, and showed a better replication for the factor AN in particular.

Second, several authors in the literature criticized the use of the two inversely scaled items since these items do not reflect direct aggression. Furthermore, these items are considered more related to aggression rejection than to acceptance of aggression characteristics (Bryant, Smith, 2001; Nakano, 2001; Ramirez et al., 2001; Vigil-Colet et al., 2005). The results of this study, similarly to a study conducted in Japan, suggested that the AQ may be improved psychometrically if the two inversely scored items were removed from the scale. Specifically in our study based on EFA and CFA, these two items had negligible loadings on their particular factors. Omission of these two items resulted in an improvement of model fit, suggesting that these items are do not provide substantial, additional information with regard to the remaining items.

Third, since Bryant and Smith (2001) reported that most of the information contained in the original 29-item scale can be reproduced reliably with a substantially shortened, 12-item version of the scale, we subjected this version to further psychometric validation based on our data. Similar to the Bryant and Smith findings, our results showed a considerable improvement in model fit over the original scale. However additional inspection of the results indicated that this increase in model fit, at least in the current study, was attributable to a singular covariance structure (Heywood Case) due to the close association between the 12 items selected for the analysis.

With regard to the external validation of the factors, we identified statistically significant gender differences with men having higher PA and VA scores, with a better separation for physical than verbal aggression. These findings are similar to those reported in many previous studies (Meesters et al., 1996; Nakano, 2001; Ramirez et al., 2001). In the original study of Buss and Perry, men had significantly higher scores on PA, VA, and HS, with the largest effect size reported for the PA factor. In our study, no significant differences between men and women were found on AN and HS, which parallels Meesters et al. and Nakano's results. In contrast to our findings, Collani and Werner (2005) found higher scores in the AN subscale for females. Consistent with our findings, all previous studies reported a gender difference in aggression, especially with regard to PA (Buss and Perry, 1992; Harris, 1995; Nakano, 2001; Ramirez et al., 2001; Fossati et al., 2003; Collani and Werner, 2005).

## References

Ando A, Soga S, Yamasaki K, Shimai S, Shimada H, Utsuki N, Oashi O, Sakai A. Development of the Japanese version of the Buss–Perry Aggression Questionnaire (BAQ). Shinrigaku Kenkyu 1999; 70: 384–392.

Beck AT, Beck RW. Shortened version of BDI. Post Grad Med 1972; 52: 81–85.

Bernstein IH, Gesn PR. On the dimensionality of the Buss/Perry Aggression Questionnaire. Behav Res Therapy 1997; 35: 563–568.

Bryant FB, Smith BD. Refining the architecture of aggression: a measurement model for the Buss-Perry Aggression Questionnaire. J Res Personality 2001; 35: 138–167.

Buss AH, Perry M. The Aggression Questionnaire. J Personality Soc Psychol 1992; 63: 452–459.

Cattell R. The Scientific Use of Factor Analysis. New York: Plenum, 1978.

Collani D, Werner R. 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 2005; 38: 1631–1643.

Cronbach L. Coefficient alpha and the internal structure of tests. Psychometrika 1951; 16: 297–334.

Felsten G, Hill V. Aggression Questionnaire hostility scale predicts anger in response to mistreatment. Behav Res Therapy 1999; 37: 87–97.

Fossati A, Maffei C, Acquarini E, Di Ceglie A. Multigroup confirmatory component and factor analyses of the Italian



DOI: 10.1002/mpr

version of the aggression questionnaire. Eur J Psychological Assessment 2003; 19: 54–65.
Gallardo D, Maydeu-Olivares A, Pueyo AA, Kramp U. A Spanish adaptation of the Aggression Questionnaire-refined: preliminary results. Barcelona: Violence Studies. Universitat Barcelona, 2004.
García-León A, Reyes GA, Vila J, Pérez N, Robles H, Ramos MM. The Aggression Questionnaire: a validation study in student samples. The Spanish Journal of Psychology 2002; 5(1): 45–53.
Harris JA. Confirmatory factor analysis of the Aggression Questionnaire. Behav Res Therapy 1995; 33: 991–993.
Harris MB, Knight-Bohnhoff K. Gender and aggression II: personal aggressiveness. Sex Roles 1996; 35: 27–42.
Hollander E, Stein D (eds). Impulsivity and Aggression. New York: John Wiley & Sons, 1995.
Kaiser HF. The application of electronic computers to factor analysis. Educational and Psychological Measurement 1960; 20: 141–151.
Loehlin JC. Latent Variable Models. Hillsdale, NJ: Lawrence Erlbaum Associates, 1987.
Lovaš L, Trenková O. Aggression and perception of an incident. Studia Psychologica 1996; 38: 265–270.
Meesters C, Muris P, Bosma H, Schouten E, Beuving S. Psychometric evaluation of the Dutch version of the Aggression Questionnaire. Behav Res Therapy 1996; 34: 839–843.
Morren M, Meesters C. Validation of the Dutch version of the aggression questionnaire in adolescent male 405 offenders. Aggress Behav 2002; 28: 87–96.
Nakano K. Psychometric evaluation on the Japanese adaptation of the aggression questionnaire. Behav Res Therapy 2001; 39: 853–858.
Palmer EJ, Thakordas V. Relationship between bullying and scores on the Buss–Perry Aggression Questionnaire among imprisoned male offenders. Aggress Behav 2005; 31: 56–66.
Ramirez JM, Andreu JM, Fujihara T. Cultural sex differences in aggression: a comparison between Japanese and Spanish students using two different inventories. Aggress Behav 2001; 27: 313–322.
Rosenberg M. Society and the Adolescent Self-image. Princeton, NJ: Princeton University Press, 1965.
Thurstone LL. Multiple-factor Analysis. Chicago, IL: University of Chicago Press, 1947.
Van Praag HM, Plutchik R, Apter A (eds). Violence and Suicidality. Perspectives in Clinical and Psychobiological Research. New York: Brunner/Mazel, 1990.
Vigil-Colet A, Lorenzo-Seva U, Codorniu-Raga MJ, Morales F. Factor structure of the Buss–Perry Aggression Questionnaire in different samples and languages. Aggress Behav 2005; 31: 601–608.
Volavka J. Neurobiology of Violence, 2nd edition. Washington, DC: American Psychiatric Publishing, 2002.

*Correspondence: József Gerevich, 1276 Budapest, PO Box 216, Hungary. Telephone 36-30-4621319.*
*Email: gerevichj@axelero.hu*



DOI: 10.1002/mpr