Elsevier Editorial System(tm) for
Personality and Individual Differences
Manuscript Draft

Manuscript Number: PAID-D-15-01131R1

Title: Short Communication The Factorial Validity and Reliability of three versions of the Aggression Questionnaire using Confirmatory Factor Analysis and Exploratory Structural Equation Modelling

Article Type: Short Communication

Section/Category: Short Communication (<3000 words)

Keywords: Exploratory Structural Equation Modelling; Confirmatory Factor Analysis; Aggression Questionnaire; Invariance Testing

Corresponding Author: Dr. Michael Thomas McKay,

Corresponding Author's Institution:

First Author: Michael Thomas McKay

Order of Authors: Michael Thomas McKay; John L Perry; Séamus A Harvey

Abstract: The Aggression Questionnaire (AQ) measures aggression in four domains: Anger, Hostility, Physical Aggression and Verbal Aggression. Moreover, a number of shorter versions of the AQ have emerged. The present study used a large sample of adolescents to test three versions of the AQ. In each case we examined a unidimensional model, a hierarchical model, and a four-factor model. Results of Confirmatory Factor Analysis revealed limited support for a unidimensional model in any of the AQ forms, with results supporting the widely-used four-factor model, and to a lesser extent, the hierarchical model. Fit indices for both short-forms of the AQ using Exploratory Structural Equation Modelling were very good. However, results also revealed only partial gender invariance for both scales.

*Title page with author details

**Short Communication**

# The Factorial Validity and Reliability of three versions of the Aggression Questionnaire using Confirmatory Factor Analysis and Exploratory Structural Equation Modelling


**Michael. T. McKay**[1]

Centre for Public Health, Liverpool John Moores University, Liverpool, UK

**John. L. Perry**

University of Hull, Hull, UK

**Séamus. A. Harvey**

Centre for Public Health, Liverpool John Moores University, Liverpool, UK

[1] Corresponding Author. Centre for Public Health, Liverpool John Moores University, 15-21 Webster Street, Liverpool, L3 2ET. Email: M.T.McKay@ljmu.ac.uk

Highlights

Three versions of the Aggression Questionnaire were examined; CFA suggested that the four-factor model was optimal; Fit indices improved with the use of ESEM in all versions; Results revealed only partial gender invariance

**Short Communication**

# The Factorial Validity and Reliability of three versions of the Aggression Questionnaire using Confirmatory Factor Analysis and Exploratory Structural Equation Modelling

## Abstract

The Aggression Questionnaire (AQ) measures aggression in four domains: Anger, Hostility, Physical Aggression and Verbal Aggression. Moreover, a number of shorter versions of the AQ have emerged. The present study used a large sample of adolescents to test three versions of the AQ. In each case we examined a unidimensional model, a hierarchical model, and a four-factor model. Results of Confirmatory Factor Analysis revealed limited support for a unidimensional model in any of the AQ forms, with results supporting the widely-used four-factor model, and to a lesser extent, the hierarchical model. Fit indices for both short-forms of the AQ using Exploratory Structural Equation Modelling were very good. However, results also revealed only partial gender invariance for both scales.

Keywords: Exploratory Structural Equation Modelling; Confirmatory Factor Analysis; Aggression Questionnaire; Invariance Testing

## 1. Introduction

The Aggression Questionnaire (AQ; Buss & Perry, 1992) consists of 29 items grouped into four factors: Physical Aggression, Verbal Aggression, Anger and Hostility. While many validation studies of the scale have been reported (e.g., Fossati, Maffei, Acquarini, & Di Ceglie, 2003), questions remain about its structure and psychometric properties. Bryant and Smith (2001) explored the factor structure of the AQ in three samples of undergraduates and reported that the four-factor model only produced a modest fit. Based on item loadings, they developed a 12-item short form of the AQ (hereafter AQ-SF) whose fit indices they reported as adequate to good. Some subsequent studies have supported this four-factor short form (e.g., Abd-El-Fattah, 2013).

More recently, another 12-item short form of the AQ, the Brief Aggression Questionnaire (BAQ), has been proposed as valid and reliable (Webster et al., 2014). These authors reported that across five studies, the BAQ showed theoretically consistent patterns of convergent and discriminant validity with other self-report measures, a four-factor structure, adequate recovery of information using item response theory methods, and adequate temporal stability and convergent validity with behavioural measures of aggression.

The present study sought to examine the properties of these three versions of the AQ in a large sample of adolescents in the United Kingdom.

## 2. Methods

### 2.1 Participants

The sample consisted of 1,004 high school students (Male = 520 [51.8%]) in school grades 9 through 12 (ages 13- to 16-years old). An 'opt out' passive consent, approved by the University Ethics Committee, ensured that parents received detailed information on the study.

### 2.2 Measure

The Aggression Questionnaire (AQ) (Buss & Perry, 1992) consists of 29 items which represent four subscales of the questionnaire: (i) Verbal Aggression (VA); (ii) Physical Aggression (PA); (iii) Anger (A); and (iv) Hostility (H). Internal consistency reliabilities reported by Buss and Perry (1992) were as follows: PA = 0.85, VA = 0.72, A = 0.83, H = 0.77, and the total score = 0.89.

### 2.3 Analyses

The dimensionality of the scales was assessed using Confirmatory Factor Analysis (CFA) and Exploratory Structural Equation Modeling (ESEM) in Mplus 7 (Muthén & Muthén, 2012) and the MLM estimator. In CFA independent cluster models (CFA-ICM), non-significant cross-loadings are constrained to zero. As such, negligible cross loadings are considered as misspecifications. ESEM (Asparouhov & Muthén, 2009) allows all observed variables to load on all latent variables. This enables freely estimated cross-loadings, has less restrictive assumptions than CFA, and potentially provides more valid estimates (Marsh, Nagengast, & Morin, 2012). As such, in psychological scales composed of indicators with many nonzero cross-loadings, ESEM is a viable alternative to CFA (Marsh et al., 2009).

An oblique geomin rotation, as recommended by Marsh et al. (2009), with an epsilon value of 0.5 and maximum likelihood estimation was used in all ESEM analyses as recommended when there are more than four response categories (e.g., Beauducel & Herzberg, 2006) and data may not be normally distributed (Bentler &

Wu, 2002). The indices used to test model fit were $\chi^2$, comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR). Although Hu and Bentler's (1999) cut-offs (i.e., >.95/.90 for CFI and TLI, <.05/.08 for RMSEA, and <.06/.08 for SRMR for good and acceptable fit respectively) are typically cited, Perry and colleagues (2015) suggested that strict adherence to these values is likely to lead to erroneous results, as factor loadings in social sciences are typically lower (see, e.g., Heene, Hilbert, Draxler, Ziegler, & Bühner, 2011). We also examined standardized parameter estimates. Factor loadings for CFA were interpreted using Comrey and Lee's (1992) recommendations (i.e., >.71 = excellent, >.63 = very good, >.55 = good, >.45 = fair and >.32 = poor). Multigroup CFA was conducted on best fitting models to examine measurement invariance across gender.

## 3. Results

Results of model fit are displayed in Table 1. The AQ demonstrated unsatisfactory model fit. Fit indices for the AQ-SF were better, in fact, the fit indices for the four-factor CFA model were borderline "good" fit. Results for the BAQ were reasonable with both relative and absolute indices achieving minimum "acceptable" thresholds for the four-factor model. In all models, the unidimensional model did not fit well. Table 2 shows that factor correlations were mostly moderate to moderately high.

**Table 1**

Table 2 also displays alpha values for all factors. These (.64 ≤ α ≤ .90) were mostly in the acceptable range for all factors of the AQ and AQ-SF, except for the alpha value for VA on the BAQ.

**Table 2**

ESEM analysis yielded fair to good model fit indices. As ESEM includes all cross-loadings, standardized parameter loadings were assessed (Table 3). The loadings of the original AQ were reasonable, although five items failed to load onto their intended factor and six items cross loaded at > .30 on a factor other than their intended. The AQ-SF encountered an identification problem due to large standard errors from item 22. Consequently, this was removed to enable identification. Of the remaining 11 items, eight loaded > .55 on their intended factor and only three items presented any statistically significant cross-loading onto another factor, two of which (.06 and .08) can be considered as negligible. The BAQ loadings were superior to the other models. Ten of the 12 items loaded > .55 on their intended factor.

**Table 3**

To examine measurement invariance across genders, we performed multigroup CFA on the AQ-SF and BAQ (Table 4). Configural invariance was assessed by replicating the CFA-ICM (independent cluster model) across males and females. Second, factors were constrained to test metric invariance. Third, we examined scalar invariance by constraining factors and item intercepts. Finally, residual variance was tested by constraining factors, item intercepts, and factor means. Model invariance is supported by little or no change in model fit on the increasingly constrained models. Cheung and Rensvold (2002) suggested ΔCFI ≤.01, although Meade, Johnson and Braddy (2008) suggest a much stricter criteria of ΔCFI ≤.002.

The initial model fit was acceptable for AQ-SF and borderline acceptable for BAQ. This remained for configural invariance. Metric invariance was supported using the more liberal criteria of ΔCFI <.01 for both scales. However, scalar and residual invariance could not be supported, suggesting that there is a gender effect in the structure of the scales. The AQ-SF presented greater invariance across gender than the BAQ did.

**Table 4**

## 4. Discussion

The present study examined the psychometric properties of three versions of the AQ, using ESEM and the more typical CFA. The results suggest that both short-forms are viable but are inconclusive in terms of which scale is optimal.

Using CFA, the AQ demonstrated unsatisfactory model fit, the AQ-SF demonstrated borderline good fit (and superior fit in comparison to the AQ and BAQ) for the four factor and hierarchical models, and the BAQ demonstrated acceptable fit for the four factor model. In contrast, using ESEM, all three scales demonstrated fair to good model fit for the four factor model, with the BAQ demonstrating best fit. Regardless of whether CFA or ESEM was employed, the findings support the four factor model of aggression in the AQ-SF and the BAQ.

In terms of loading onto their hypothesised factors, problems were evident with the H items, where four loaded substantively more strongly onto the VA factor (.33 $\leq$ loading $\leq$ .64) than H (.01 $\leq$ loading $\leq$ .17). Another H item loaded .30 on its hypothesised factor, with a substantive loading on all other factors (.18 $\leq$ loading $\leq$ .28). Another H item loaded only .14 on H but .44 on Anger. This suggests particular problems with the H construct in this sample. There were also a number of issues with two VA items and one PA item substantively cross loading.

In order for gender difference results to be considered valid and reliable, gender invariance must be assured. Configural invariance was supported in both AQ short forms, indicating that males and females held the same basic perception of trait aggression, distinguished between the four factors, and identified the same items which loaded onto these four factors. Metric invariance was also supported, indicating that males and females used comparable conceptual frames of reference when responding to the items. However, scalar invariance was not supported, indicating that males and females did not use the response scale in a comparable way. Further, residual variance was not supported, indicating that the differences between male and female responses to items were not accounted for by their differences on the four factors. These findings demonstrate that there is a gender effect in the structure of the scales, with the BAQ demonstrating greater gender variance.

However, the results of the present study should be interpreted in the context of problems assessing '*short forms*' of scales administered in their original longer format (Knowles & Condon, 2000), in particular the fact that responses to items on scales "often involves more than responding to the semantic content of the item. Respondents interpret the items within a context. As the context for an item changes, even as its position in the test changes, the meaning of the item may shift" (p.250). In conclusion, the present study offers some support for the psychometric properties of the four factor AQ-SF and BAQ. The results underline the importance of on-going and rigorous assessment of scale properties, in particular when assessing variables sensitive to cultural influence like aggression.

## References


Abd-El-Fattah, S.M. (2013). A cross-cultural examination of the Aggression Questionnaire–Short Form among Egyptian and Omani Adolescents. *Journal of Personality Assessment*, *95*, 539-548.

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

Beauducel, A., & Herzberg, P.Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, *13*, 186–203.

Bentler, P.M., & Wu, E.J.C. (2002). EQS 6 for Windows User's Guide. Multivariate Software.

Bryant, F.B., & Smith, B.D. (2001). Refining the architecture of aggression: A measurement model for the Buss-Perry aggression questionnaire. *Journal of Research in Personality*, *35*, 138-167.

Buss, A.H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, *63*, 452-459.

Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for measurement invariance. *Structural Equation Modeling*, *9*, 233-255.

Comrey, A.L., & Lee, H.B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Lawrence Erlbaum.

Fossati, A., Maffei, C., Acquarini, E., & Di Ceglie, A. (2003). Multi-group confirmatory component and factor analyses of the Italian version of the Aggression Questionnaire. *European Journal of Psychological Assessment*, *19*, 54-65.

Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods*, *16*, 319–336.

Hu, L., & Bentler, P.M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling*, *6*, 1−55.

Knowles, E.S., & Condon, C.A. (2000). Does the rose still smell sweet? Item variability across test forms and revisions. *Psychological Assessment*, *12*, 245-252.

Marsh, H.W., Muthén, B., Asparouhov, A., Lüdtke, O., Robitzsch, A., Morin, A.J.S., et al. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, *16*, 439–476.

Marsh, H.W., Nagengast, B., & Morin, A.J.S. (2012). Measurement invariance of Big-Five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and La Dolce Vita effects. *Developmental Psychology*

Meade, A.W., Johnson, E.C., & Braddy, P.W. (2008). Power and sensitivity of alternate fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*, 568-592.

Muthén, B.O., & Muthén, L.K. (2012). Mplus version 7 [Computer Programme]. Los Angeles, CA: Muthén & Muthén.

Perry, J.L., Nicholls, A.R., Clough, P.J., & Crust, L. (2015). Assessing model fit: Caveats and recommendations for confirmatory factor analysis and exploratory structural equation modeling. *Measurement in Physical Education and Exercise Science, 19*, 12-21.

Webster, G.D., De Wall, C.N., Pond Jr., R.S., Deckman, T., Jonason, P.K., Le, B.M., et al. (2012). The Brief Aggression Questionnaire; psychometric and behavioral evidence for an efficient measure of trait aggression. *Aggressive Behavior*, *40*, 120-139.

Table(s)

**Table 1**

***CFA and ESEM model fits for each model***

| Model | $\chi^2$ | *df* | CFI | TLI | SRMR | RMSEA (90% CI) |
|---|---|---|---|---|---|---|
| *CFA* | | | | | | |
| AQ, unidimensional | 3708.15* | 377 | .667 | .642 | .088 | .094 (.091, .097) |
| AQ, 4-factor | 1973.27* | 371 | .840 | .825 | .070 | .066 (.063, .069) |
| AQ, hierarchical | 2020.83* | 373 | .835 | .821 | .074 | .066 (.064, .069) |
| AQ-SF, unidimensional | 677.67* | 54 | .709 | .644 | .077 | .107 (.100, .115) |
| AQ-SF, 4-factor | 243.89* | 48 | .936 | .912 | .064 | .064 (.056, .072) |
| AQ-SF, hierarchical | 255.77* | 50 | .933 | .911 | .066 | .064 (.056, .072) |
| BAQ, unidimensional | 904.59* | 54 | .695 | .627 | .088 | .125 (.118, .133) |
| BAQ, 4-factor | 309.43* | 48 | .906 | .871 | .070 | .074 (.066, .082) |
| BAQ, hierarchical | 366.29* | 50 | .887 | .850 | .071 | .079 (.072, .087) |
| *ESEM* | | | | | | |
| AQ, 4-factor | 933.04* | 296 | .936 | .913 | .031 | .046 (.043, .050) |
| AQ-SF, 4-factor[a] | 17.19 | 17 | 1.00 | 1.00 | .009 | .003 (.000, .029) |
| BAQ, 4-factor | 21.45 | 24 | 1.00 | 1.03 | .009 | .000 (.000, .022) |

[a]Item 22 created an identification error making the model inadmissible and was therefore removed.

*Statistically significant at $p < .001$.

**Table 2**

***Factor correlations for 4-factor CFA and ESEM models***

| Scale | A | H | VA | PA |
|---|---|---|---|---|
| *AQ* | | | | |
| Anger | (.86) | .33** | .43** | .56** |
| Hostility | .53** | (.75) | .41** | .12** |
| Verbal Aggression | .71** | .59** | (.67) | .21** |
| Physical Aggression | .69** | .35** | .72** | (.90) |
| *AQ-SF* | | | | |
| Anger | (.66) | .36** | .57** | .47** |
| Hostility | .42** | (.69) | .38** | .20** |
| Verbal Aggression | .66** | .48** | (.64) | .54** |
| Physical Aggression | .64** | .31** | .65** | (.82) |
| *BAQ* | | | | |
| Anger | (.78) | .34** | .18* | .51** |
| Hostility | .50** | (.53) | -.03 | .18** |
| Verbal Aggression | .44** | .26** | (.45) | .34** |
| Physical Aggression | .57** | .36** | .66** | (.82) |

*Note*. CFA factor correlations appear below the diagonal, ESEM factor correlations appear above the diagonal. Cronbach's alpha internal consistency estimates are shown in parentheses. *Statistically significant at $p < .05$, **$p < .01$.

**Table 3**

***ESEM factor loadings for each 4-factor model***

| Item | Physical Aggression | | | Verbal Aggression | | | Anger | | | Hostility | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | M1 | M2 | M3 | M1 | M2 | M3 | M1 | M2 | M3 | M1 | M2 | M3 |
| 7 | .03 | .18* | .05 | -.11 | .12 | .03 | **.88**** | **.61**** | **.84**** | .01 | -.03 | -.01 |
| 17 | .00 | - | .03 | -.14 | - | -.02 | **.76**** | **-** | **.72**** | .03 | - | -.02 |
| 24 | -.19 | -.08 | -.07 | .08 | -.02 | .00 | **.74**** | **.78**** | **.59**** | .03 | .03 | .20** |
| 1 | .01 | .10 | - | -.11 | .04 | - | **.44**** | **.32**** | **-** | .03 | -.02 | - |
| 3 | .05 | - | - | -.10 | - | - | **.59**** | - | **-** | -.02 | - | - |
| 25 | .09* | - | - | .10 | - | - | **.72**** | - | **-** | -.16 | - | - |
| 28 | .05 | - | - | .20** | - | - | **.64**** | **-** | **-** | .01 | - | - |
| 12 | .09* | .06* | .10 | .01 | -.03 | -.07 | .01 | .04 | .14* | **.83**** | **.83**** | **.40**** |
| 26 | .25** | - | .25** | .42** | - | .08 | -.06 | - | -.04 | **.01** | **-** | **.35**** |
| 27 | -.16 | - | -.05 | .64** | - | -.03 | .04 | - | -.02 | **.17**** | **-** | **.77**** |
| 6 | -.25 | **-** | - | .18** | - | - | .28** | - | - | **.30**** | - | **-** |
| 9 | .04 | -.04 | - | .03 | .10 | - | .00 | -.03 | - | **.79**** | **.79**** | **-** |
| 14 | .-.04 | .02 | - | .13* | -.02 | - | .44** | .51** | - | **.14**** | **.16**** | **-** |
| 19 | .01 | - | - | .68** | - | - | -.04 | - | - | **.07** | - | **-** |
| 21 | .08 | - | - | .33** | - | - | -.02 | - | - | **-.03** | - | **-** |
| 2 | .20** | **-** | -.01 | **.02** | - | **.67**** | .05 | - | -.02 | -.17 | - | -.09 |
| 11 | .44** | - | .33** | **.08** | **-** | **.45**** | .13* | - | .06 | -.07 | - | .05* |
| 20 | .17** | .02 | .19** | **.43**** | **.52**** | **.09** | .13* | .02 | .11* | -.06 | -.01 | .30** |
| 5 | .14** | -.03 | - | **.33**** | **.56**** | **-** | .13* | .01 | - | .08 | .08 | - |
| 16 | .26** | .06 | - | **.27**** | **.67**** | **-** | .23** | .02 | - | .04 | .00 | - |
| 10 | **.76**** | **.77**** | **.75**** | -.01 | .05 | .00 | .09* | .05 | .11** | .04 | -.01 | -.01 |
| 15 | **.75**** | - | **.75**** | .10* | - | .10 | -.03 | - | -.04 | -.06 | - | .00 |
| 18 | **.79**** | **80**** | **.82**** | .02 | -.02 | -.09 | -.01 | -.02 | .00 | .05 | .04 | -.01 |
| 4 | **.60**** | - | **-** | -.06 | - | - | .24** | - | - | .05 | - | - |
| 8 | **.74**** | **-** | **-** | -.04 | - | - | -.01 | - | - | .06 | - | - |
| 13 | **.48**** | - | **-** | .01 | **-** | - | .21** | - | - | .02 | - | - |
| 22 | **.58**** | **inad** | **-** | .16** | Inad | - | .15** | inad | - | -.01 | inad | - |
| 23 | **.47**** | - | **-** | -.05 | - | - | .01 | - | - | .06 | - | - |
| 29 | **.37**** | - | **-** | .03 | - | - | .44** | - | - | -.03 | - | - |

*Note*. M1 = AQ, M2 = AQ-SF, M3 = BAQ. Loadings on intended factors are highlighted in bold. Item 22 was inadmissible and removed from the model. *Statistically significant at $p < .05$, $^{**}p < .01$.

**Table 4**

***Measurement invariance for AQ-SF and BAQ four factor models***

| Model | $\chi^2$ | *df* | $\Delta \chi^2$ | $\Delta df$ | CFI | TLI | SRMR | RMSEA (90% CI) |
|---|---|---|---|---|---|---|---|---|
| *AQ-SF* | | | | | | | | |
| Configural invariance | 291.21* | 96 | - | - | .934 | .910 | .067 | .064 (.055, .072) |
| Metric invariance | 305.28* | 104 | 14.07 | 8 | .932 | .914 | .071 | .062 (.054, .070) |
| Scalar invariance | 337.32* | 112 | 32.04 | 8 | .924 | .911 | .072 | .063 (.056, .071) |
| Residual invariance | 524.97* | 116 | 187.65 | 4 | .863 | .844 | .103 | .084 (.077, .091) |
| *BAQ* | | | | | | | | |
| Configural invariance | 349.41* | 96 | - | - | .904 | .868 | .074 | .073 (.064, .081) |
| Metric invariance | 363.52* | 104 | 14.11 | 8 | .901 | .875 | .076 | .071 (.063, .079) |
| Scalar invariance | 455.99* | 112 | 92.47 | 8 | .869 | .846 | .078 | .078 (.071, .086) |
| Residual invariance | 700.08* | 116 | 244.09 | 4 | .778 | .748 | .115 | .100 (.093, .107) |

*Statistically significant at $p < .001$.