Personality and Individual Differences 90 (2016) 12–15

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

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 [a,*], John L. Perry [b], Séamus A. Harvey [a]

[a] Centre for Public Health, Liverpool John Moores University, Liverpool, UK
[b] University of Hull, Hull, UK

ARTICLE INFO

Article history:
Received 31 July 2015
Received in revised form 8 October 2015
Accepted 10 October 2015
Available online 20 October 2015

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

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.

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 1004 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 of Liverpool 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 Modelling

* Corresponding author at: Centre for Public Health, Liverpool John Moores University, 15-21 Webster Street, Liverpool L3 2ET, UK.
E-mail address: M.T.McKay@ljmu.ac.uk (M.T. McKay).

http://dx.doi.org/10.1016/j.paid.2015.10.028
0191-8869/

(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 χ^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, Nicholls, Clough, and Crust (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 2 also displays alpha values for all factors. These ($.64 \le \alpha \le .90$) 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.

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.

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 $\Delta CFI \le .01$, although Meade, Johnson, and Braddy (2008) suggest a much stricter criterion of $\Delta CFI \le .002$.

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 $\Delta CFI < .01$ 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 1
CFA and ESEM model fits for each model.

Model	χ^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 $p < .05$.
** Statistically significant at $p < .01$.

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 Hostility items, where four loaded substantively