

Original article

## Assessing aggressiveness quickly and efficiently: the Spanish adaptation of Aggression Questionnaire-Refined version

David Gallardo-Pujol \*, Uwe Kramp, Carlos García-Forero, Meritxell Pérez-Ramírez,  
Antonio Andrés-Pueyo

*Department of Personality, Faculty of Psychology, University of Barcelona, Passeig de la Vall d'Hebron 171, 08035 Barcelona, Spain*

Received 3 June 2005; accepted 13 February 2006

Available online 03 May 2006

### Abstract

The assessment of aggressiveness and the prediction of aggression has become a relevant research and applied topic in Psychiatry and Psychology. There have been many attempts in order to get a fast and reliable tool to measure aggression. Buss and Durkee started the pathway, and recently Bryant and Smith developed a tool with an enormous potential, a fast-applicable, reliable and valid test. We herein report a Spanish adaptation of this test and we show that aggressiveness can be measured rapidly, and in a simple, valid and reliable way across different populations. We focus on the discriminant capacity of this test to detect aggressive individuals.

© 2006 Elsevier Masson SAS. All rights reserved.

*Keywords:* Aggression; Personality; Impulsivity; Questionnaires

### 1. Introduction

Aggression and aggressiveness are psychological topics of particular relevance in forensic and clinical contexts. To predict if somebody will behave aggressively is important within the clinical environment [29,16,27,17]. But it is not only important for the patient himself; in some cases it can be useful for the staff as well [31] to predict the behaviour of the patients within the institutions. At the prison or the court, often the judgement of a specialist is required in order to assess the premeditated nature of an offence [3] or the presence or absence of psychopathological disorders in the offender, i.e. psychopathy [5]. In many cases, the violence risk assessment does not include self-report measures of aggression. Basically this is because in these contexts, the nature of the assessed individual often restricts the use of this type of measures, due to various factors (reading limitations, time restrictions, etc.), so the existence of

a short and easily-applicable self-reported measure would help those professionals who work in the fields mentioned above.

In order to achieve a good self-report measure of aggression, many researchers have developed their own questionnaires. In 1957, Buss and Durkee [10] developed the Buss–Durkee Hostility Inventory, a world-wide questionnaire which is still used. This test has become one of the required tools in any research on aggression, in spite of its psychometric limitations. Based on the BDHI, Buss and Perry [11] redefined it to improve its psychometrical properties, and the result was the Aggression Questionnaire (AQ), which measures four aspects of aggression: Physical Aggression and Verbal Aggression, which involve hurting or harming others, represent the instrumental or motor component of behaviour; Hostility, which consists of feelings of ill-will and injustice, represents the cognitive component of behaviour; and Anger, which involves physiological arousal and preparation for aggression, represents the emotional or affective component of behaviour. Incidentally, this questionnaire has been used to investigate topics such as suicidal ideation [23] and eating disorders [24], just to cite some examples, even in violence risk assessment, explor-

\* Corresponding author.

*E-mail address:* [david.gallardo@ub.edu](mailto:david.gallardo@ub.edu) (D. Gallardo-Pujol).

ing the relationships between this questionnaire and the PCL-R [28].

In 2001, Bryant and Smith [9] had taken again about this topic and they developed a shorter form of the AQ. This shorter form comprises 12 item-stems from Buss and Perry's AQ [11], less than a half of the original 29-items test. This improvement has made possible to have a short and reliable test available in order to use it in critical contexts such as those mentioned above, in which the application conditions are often challenging. In addition, the few number of items of this test makes possible to include them within a larger battery of items.

Within the Spanish-speaking context, to our knowledge, few studies have been conducted using AQ [1,15] and, any using its refined form [9]. Recently, Vigil-Colet et al. [36,25] developed a reduced form comparing the AQ factor structure across different samples, and selecting the items with a high congruence among samples, so they obtained a 20-items long AQ. Nevertheless, their aim was not to search specifically for the shortest form of AQ; rather, their aim was to find a set of items with a simple factor structure that was stable across cultural differences, that is, reasonably free from cultural differences. Thus, there is a lack of an even more rapidly applicable and reliable tool in the Spanish-speaking context, a tool not only valid but also useful for the professional, in the sense that it provides information about prognosis, treatments and social correlates [18]. So, our aim was to adapt the shortest form of the AQ to the Spanish-speaking context and fill this gap. Doing so, we would facilitate and promote research with Spanish-speaking participants as well.

## 2. Material and methods

### 2.1. Participants

In order to adapt the short form of the AQ [9] to Spanish, three subsamples were considered:

First, a huge group of 1047 volunteers. The mean age was 34.34 (S.D. = 14.30), ranging from 16 to 75. With respect to their educational levels, a 39.50% of the sample achieved a graduate level, a 16.50% has a secondary school level, a 15.80% has an elementary school level, a 14% has no studies, and finally in a 14.2% of this subsample these data were not available. Regarding their job status, a 25% of the samples were students, a 42.5% has a regular job, a 6.7% of the subsample does not work, and a 25.8% of the data are missing. This subsample was composed by a 46.6% of women and a 45.4% of men, 8% are missing. This subsample was informed that these data would be used for research purposes, and that the participation was absolutely voluntary, they agree if they return the questionnaires.

Second, a subsample group of 81 students from the Police School of Catalonia. The mean age was 26.48 (S.D. = 3.98), ranging from 21 to 40. With respect to their educational levels, a 14.8% of the sample achieved a graduate level, a 63% has a secondary school level, and a 22.2% has an elementary school level. Regarding their job status, all of them were students of

Police. This subsample was composed by a 14.8% of women and an 84% of men, 1.2% are missing. The participants of this subsample were informed that part of the data from their global assessment would be used for research purposes. Those who did not agree were not included in the analysis.

And third, the last one, a subsample of 140 offenders throughout three prisons of the penitentiary system of the Autonomous Catalan Government. The mean age was 36.06 (S.D. = 10.17), ranging from 21 to 68. With respect to their educational level, a 6.4% of the sample achieved a graduate level; a 19.3% has a secondary school level, a 53.6% has an elementary school level, and a 20.7% has no studies. All of them were inmates. This subsample was composed by a 19.3% of women and an 80.7% of men. The participants of this subsample were informed that part of the data from their global assessment would be used for research purposes. Those who did not agree were not included in the analysis.

In all three samples, the assessment was conducted under conditions of strictly anonymity, and the respondents were asked to be the more sincere they can. In addition, this study is part of a global study approved by the Ethics' Committee of the University of Barcelona.

There was no missing data due to no-response to items. Even so, some of the respondents were removed from the sample due to their influence on multivariate normality and anomalies in their responses, such as outlier cases in total scores.

### 2.2. Instrument

For this study, it was used the refined version of the AQ [9], which consists of 12 Likert-type items rated on a 6-point scale. The AQ is organised into four scales: Physical Aggression or PA (three items), Verbal Aggression or VA (three items), Anger or ANG (three items) and Hostility or HO (three items). Our version consisted of 12 Likert-type items rated on a 5-point scale ranging from "Never" to "Always". The authors restored the scale's mid-point which was eliminated in the English version by Bryant and Smith [9]. This modification was done taking into account the following criteria: (a) the traditional Likert-type response format (five response alternatives with a neutral point [20,21]) seems to be more adequate in the context of cross-cultural comparisons [30]; (b) empirical support exists in favour to the traditional Likert-type format, in front of a forced-choice format [35]; (c) the reliability of a personality measurement instrument does not vary significantly when is reduced only one response category [19]; and (d) because the original version of AQ [11] has a 5-point scale.

The adaptation of the refined version of AQ was done using the back-translation method, which is a judgmental procedure for investigating the conceptual equivalence (i.e. symmetry) of the original and translated versions, necessary for valid cross-cultural comparisons [4]. Thereby, a bilingual translator translated the American Questionnaire into Spanish. Secondly, another bilingual translator translated back the Spanish AQ into English. Finally, the original source and the back-translated

items were compared for non-equivalence of meaning by an expert, and any discrepancies were noted. This iterative process of translation and back-translation was continued until no semantic differences were noticed between both questionnaire forms [6]<sup>1</sup>.

Impulsivity was assessed using the Spanish translation of Luengo et al. [22] of Barratt's Impulsiveness Scale (BIS-10: [2]). The BIS-10 has been widely used among clinicians and researchers and it consists of three subscales: Motor Impulsiveness (Im, 11 items), Cognitive Impulsiveness (Ic, 11 items) and Non-planning Impulsiveness (Inp, 12 items). It has a total score as well, which indicates a global measure of impulsiveness. Respondents may answer by means of a Likert-like scale of four options from "never" to "always", but without the mean item. BIS-10 was chosen to assess the concurrent validity due to the close relationship that exists between impulsiveness and aggressiveness (see Barratt and Slaughter [3] for a review).

### 2.3. Procedure

Firstly, the test was administered to university students within the classroom time. Then, the subjects were given instructions to collect the remaining data, by using the "snowballing approach" in which students passed on the test to their relatives and friends. They were asked to recruit two or three subjects from both sexes, and were trained on how to administer the questionnaire. In the second group, the test was administered included in the formal process of candidates' selection at the Police School of Catalonia. The test was included within a battery of tests used to assess the individuals in their process to become a policeman. Finally, the third group was collected throughout three prisons of Catalonia, as mentioned above. The test was administered individually by the prison's psychologists to inmates. To calculate test-retest reliability, the test was administered to the control sample twice, with a 4 months minimum interval.

### 2.4. Overview of analyses

Our analyses are aimed to five purposes: first, item analysis of the scales, including differential item functioning (DIF) detection; second, a multivariate MANOVA conducted to estimate differences between sex and population, which is also an indicator of construct validity as well; third, internal consistency of the four factors, evaluated by the alpha coefficient; four, correlations among scales to assess construct validity

<sup>1</sup> No permission is needed if the questionnaire is used for research purposes. Users can format their own printed versions. The score for each scale is the sum of the ratings for its items. The total score for aggressiveness is the sum of these scale scores. The respondents were given to the following instructions: "Here are a number of characteristics that may or may not apply to you. Please check the appropriate option to indicate the extent to which the phrase can be applied to you".

(see below); finally, a confirmatory factor analysis to replicate the structure of the original version.

Regarding the confirmatory factor analysis, the model fitted in this study was estimated via maximum likelihood estimation without assuming normality, as implemented in MPLUS [26] using covariances as sample statistics. To assess the degree of lack of fit of the model being estimated we shall employ the root mean squared error of approximation (RMSEA) [32], as a measure of discrepancy due to the approximation per degree of freedom. A RMSEA of about 0.05 reflects a close fit of the model in relation to its degrees of freedom, whereas a value of 0.08 or less reflects a reasonable error of approximation [7,8]. In addition, Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) were also calculated. To show a good fit of the model to the data, these indexes must be between 0.9 and 1.

In order to assess the construct validity, the police students group and the offenders group were used to estimate the possible differences between groups, in theory, from a sample of less (police) to more (offenders) aggressive than the control sample (unselected sample). It was done by means of a multivariate MANOVA.

## 3. Results

First, item analysis was conducted in order to check the item-scales properties. Table 1 shows some descriptive statistics to get a flavour of item properties in general population, which was the sample of full-scale analysis. Almost every item exhibits positive skewness, the first two categories of the scale attracting more responses while the fifth point of the ordered scales never attracts more than a 6.5% of the responses for any item. This floor effect is common in items designed to elicit extreme responses, meaning that their behaviour correspond to items that stand for high levels on the underlying construct, and that receive, accordingly, few responses on a extreme category. Despite of being extreme, all of them show good item-test correlations both for its scale and the overall total.

One more item analysis was conducted to check if differences in total scores could be due to DIF with respect to gender or population. As suggested by Swaminathan and Rogers [33], logistic regression was used, after dichotomising item responses, recodifying values 1 through 2 as 0 and 3 through 5 as 1. Wald statistic for sex ranged from 0.452 ( $P = 0.50$ ) to 5.72 ( $P = 0.02$ ), for Group showed values from 0.05 ( $P = 0.82$ ) to 5.46 ( $P = 0.02$ ). Thus, at a 0.01 nominal level, no item showed DIF with respect to either of the variables.

Second, a MANOVA was conducted in order to determine the effect of sex and population in subscale means. As equality between covariance matrices was rejected a robust  $F$  statistic was selected for contrasting mean equality between groups of sex and population. A significant effect was found, both for sex ( $F = 4.03$ ;  $df1 = 4$ ,  $df2 = 969$ ;  $P = 0.003$ ) and group ( $F = 26.64$ ;  $df1 = 8$ ,  $df2 = 1944$ ;  $P < 0.001$ ). Table 2 shows the between subjects effects of the influence of each of the variables on subscales. Interactions between sex are group are not

Table 1  
Item descriptives

Item	Subscale <sup>a</sup>	Category frequency (%)					Item descriptives				Item-total corrected correlation	
		1	2	3	4	5	Mean	S.D.	Median	Skewness <sup>b</sup>	Subscale	Scale
I1	VA	1.8	14.9	72.6	9.9	0.7	2.93	0.588	3.00	-0.310	0.33	0.341
I2	HOST	31.7	38.4	25.7	3.4	0.8	2.03	0.886	2.00	0.522	0.47	0.340
I3	PA	70.6	23.1	6.1	0.2	0.0	1.36	0.605	1.00	1.538	0.41	0.315
I4	HOST	48.9	29.4	17.8	3.4	0.5	1.77	0.892	2.00	0.917	0.51	0.443
I5	ANG	26.8	32.9	33.2	4.8	2.3	2.23	0.974	2.00	0.442	0.51	0.553
I6	VA	32.3	23.2	32.2	9.4	2.8	2.27	1.096	2.00	0.388	0.45	0.335
I7	ANG	11.6	25.2	38.5	18.2	6.5	2.83	1.064	3.00	0.093	0.40	0.381
I8	PA	63.2	28.9	5.3	2.1	0.5	1.48	0.734	1.00	1.787	0.52	0.334
I9	VA	16.2	30.3	35.1	14.2	4.2	2.60	1.050	3.00	0.233	0.47	0.437
I10	HOST	26.0	36.7	29.8	5.7	1.8	2.21	0.953	2.00	0.481	0.47	0.346
I11	PA	61.6	27.1	9.7	1.1	0.5	1.52	0.753	1.00	1.485	0.61	0.458
I12	ANG	47.8	32.2	17.2	2.3	0.5	1.75	0.852	2.00	0.905	0.47	0.467

N = 826.

<sup>a</sup> VA = Verbal Aggression; PA = Physical Aggression; HOST = Hostility; ANG = Anger.

<sup>b</sup> Skewness standard error = 0.085.

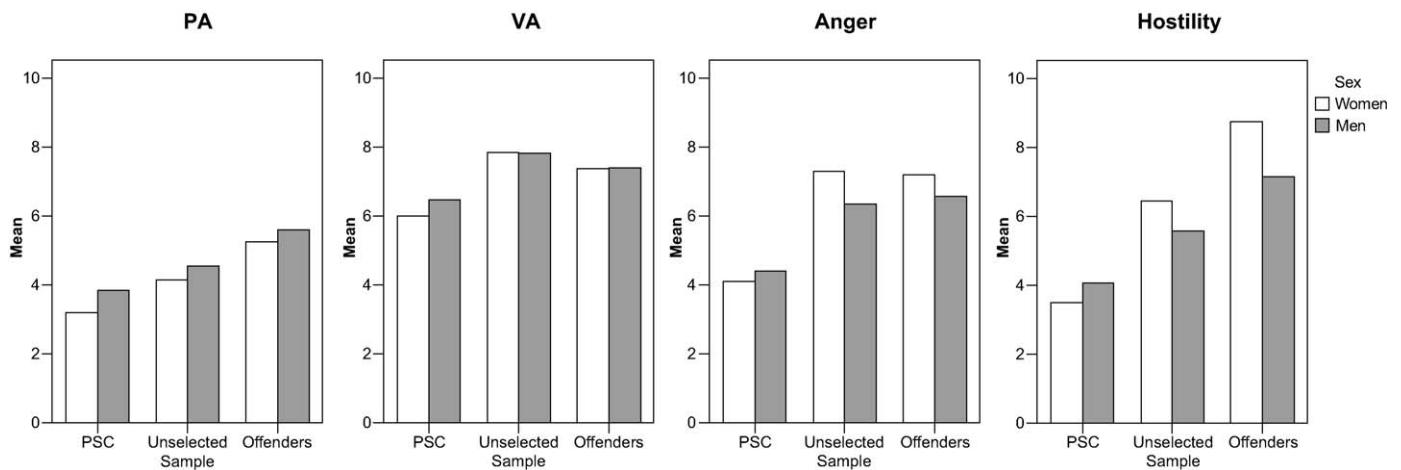


Fig. 1. Scale means by sex and population subgroups.

Notes: PA, AQ-R Physical Aggression; VA, AQ-R Verbal Aggression; PSC, Police School of Catalonia.

Table 2  
Principal effects' contrasts on AQ-R subscales

		F	df1	df2	P	Observed power (with $\alpha = 05$ )
Sex	PA	4.047	1	972	0.045	0.95
	VA	0.314	1	972	0.575	0.05
	ANG	2.117	1	972	0.045	> 0.99
	HOST	4.845	1	972	0.028	> 0.99
Population group	PA	19.080	2	972	0.000	> 0.99
	VA	11.778	2	972	0.000	> 0.99
	ANG	19.080	2	972	0.000	> 0.99
	HOST	54.675	2	972	0.000	> 0.99

Notes: PA, AQ-R Physical Aggression; VA, AQ-R Verbal Aggression, ANG: AQ-R Anger; HOST, AQ-R Hostility.

significant ( $F = 1.18$ ,  $df1 = 8$ ,  $df2 = 1940$ ,  $P = 0.31$ ). Fig. 1 shows the simultaneous effects of both population and sex.

Thus, as can be seen from Table 2 and Fig. 1, men differ from women on subscales PA, ANG and HOS. Fig. 1 shows that, on the average, men have a superior mean than women on subscale PA. This situation is inverted on subscales ANG and HOS, where women obtain greater means than men. To ana-

lyse the mean trends with respect to population, multiple comparisons with Bonferroni's adjustment were conducted with a 0.05 significance level. Comparisons show that individuals from the Police School of Catalonia's means are significantly inferior to those of the other groups. On the other hand, and with the same significance level, unselected sample means are significantly inferior to those of offenders on subscales Physi-

Table 3  
Descriptives of AQ-R subscales and total score

	Sex				Population						Total reliability (unselected sample)	
	Women		Men		PSC		Unselected sample		Offenders		$\alpha$	Test-retest
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.		
PA	4.20	1.64	4.63	1.80	3.80	1.11	4.35	1.65	5.53	2.21	0.70	0.58
VA	7.78	2.04	7.58	2.03	6.42	1.57	7.80	2.08	7.39	1.69	0.58	0.66
ANG	7.21	2.22	6.14	2.18	4.37	1.11	6.81	2.22	6.71	2.23	0.66	0.46
HOST	6.50	2.31	5.65	2.04	4.01	1.13	6.01	2.12	7.50	2.26	0.68	0.67
TOTAL	25.69	5.51	24.00	5.93	18.59	3.57	24.97	5.54	27.14	6.14	0.78	0.71

Notes: This table shows the descriptives of AQ-R subscales for population and sex, as well as reliability indices. PSC: Police School of Catalonia; PA: AQ-R Physical Aggression; VA: AQ-R Verbal Aggression; ANG: AQ-R Anger; HOST: AQ-R Hostility; TOTAL: AQ-R Total score.

cal Aggression and Anger, and no significant differences were found on subscales Verbal Aggression and Anger. These trends can be seen in Fig. 1.

Table 3 shows the descriptive statistics of the AQ-R by sex and population. In addition, it contains the statistics for the internal consistency as well as for temporal stability. Thus, with regard to the subscales, alpha coefficients range from 0.58 to 0.70, and test–retest stability index range from 0.46 to 0.67. For the whole AQ-R was found to be 0.78 for alpha coefficient and 0.71 for temporal stability index. So, the results regarding alpha coefficient suggest that AQ-R subscales have an appropriate internal consistency level, despite of the fact that they contain just three items. These alpha reliabilities are comparable to those of the original version.

Table 4 shows the correlations between AQ-R subscales and BIS-10 subscales. Despite all correlations were significant ( $P < 0.01$ ), the correlations between BIS-10 total score and AQ-R subscales ranged from 0.23 to 0.39. There is a similar effect in correlations between AQ-R total score and BIS-10 subscales, which ranged from 0.23 to 0.50. Actually, differences between AQ-R and BIS-10 were expected, as long as both tests measure two different constructs, although closely related, and thus diverging in their measures.

To evaluate factorial construct validity, the four factor model of Bryant and Smith [9] was fitted to the unselected sample. The model is an independent clusters solution in which each item is loaded by only one factor and all factors are inter-correlated. The model shows a fair fit ( $\chi^2 = 158.38$ ; degrees of freedom = 48; RMSEA = 0.053, 95% CI = 0.040 ÷ 0.060). In Fig. 2 we provide statistics for the standardised loading factors

Table 4  
Correlations between AQ-R and BIS-10 subscales

	1	2	3	4	5	6	7	8	9
1 PA	1								
2 VA	0.26	1							
3 ANG	0.34	0.44	1						
4 HOST	0.30	0.21	0.39	1					
5 TOTAL	0.64	0.68	0.79	0.69	1				
6 BIS-NPI	0.13	0.13	0.20	0.18	0.23	1			
7 BIS-MI	0.29	0.31	0.47	0.32	0.50	0.41	1		
8 BIS-CI	0.26	0.09	0.27	0.18	0.28	0.45	0.45	1	
9 BIS-TOTAL	0.29	0.23	0.39	0.31	0.43	0.83	0.79	0.74	1

Notes: All correlation coefficients are significant ( $P < 0.01$ ). AQ-R Physical Aggression; VA: AQ-R Verbal Aggression; ANG: AQ-R Anger; HOST: AQ-R Hostility; TOTAL: AQ-R Total score; BIS-NPI: BIS-10 Non-planning Impulsivity; BIS-MI: BIS-10 Motor Impulsivity; BIS-CI: BIS-10 Cognitive Impulsivity; BIS-TOTAL: BIS-10 Total score.

in each sample, as well as correlations between factors. The CFI and TLI indexes also show enough goodness-of-fit (0.93 and 0.91, respectively), as well as they are close to 1, which means a good fit of the data to the model. The complete list of the items is reported in Appendix 1.

#### 4. Discussion

The aim of our study was to adapt the AQ-refined version to the Spanish-speaking context. In the light of our results, this test seems to be a good tool in the assessment of aggressiveness, and therefore of violence. The correlations among subscales and with external scales of impulsivity, the reliability coefficients, and the good fitting of the four factor model provide support to the AQ-R’s construct validity.

As expected, all subscales of AQ-R correlated with all BIS-10 subscales, especially with Motor Impulsiveness. The correlations that we have obtained are quite similar to those obtained by Buss and Perry [11] in their original report, when they correlated AQ subscales with the Impulsiveness scale of EASI [12]. Some other authors have reported correlations between AQ subscales and traits related to impulsiveness. For instance, Tremblay and Ewart [34] report an inverse and moderate correlation between Physical Aggression and Conscientiousness measured by means of NEO-PI-R. So, our findings are consistent with previous literature.

If we compare this refined version to the reduced one obtained by Vigil-Colet et al. [36], we can see that the majority of items in our AQ-R are stable across cultures. Both versions only differ in three items (8, 12, and 25 in the original AQ from Buss and Perry). Two of them (8 and 12) belong to the Hostility factor, and the other to the Anger factor. These items may be related to a factor measuring suspiciousness, rather than hostility [36]. On the other hand, the rest of the items are shared between both versions.

Regarding the internal consistency, our results are slightly lower than those reported by Andreu-Rodríguez et al. and García-León et al. [1,15] which is consistent with the definition of “alpha” given by Cronbach [14], in which, the higher the number the items in a scale, the higher the alpha coefficient. In the same way, compared to the internal consistencies of three samples using the English refined, our version is equivalent to the results reported by Bryant and Smith [9], so, these results suggest appropriate internal consistencies in relation to the length

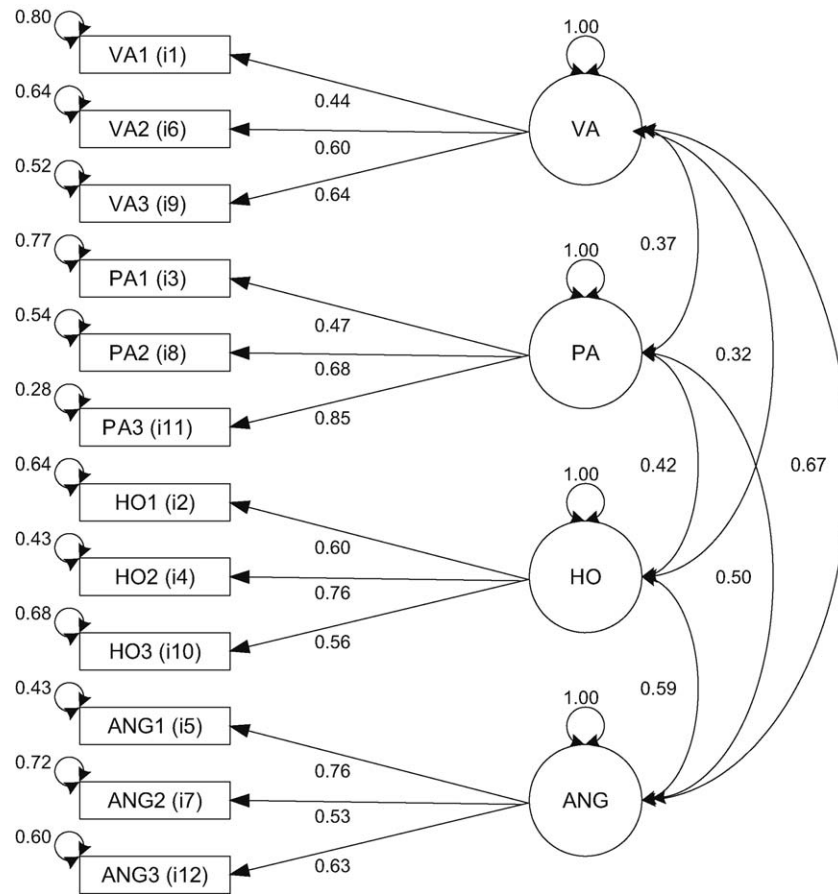


Fig. 2. Path diagram of the confirmatory factor analysis of the Spanish version of AQ-R.  
 Notes: Numbers in parenthesis indicate the corresponding item in the questionnaire. PA, AQ-R Physical Aggression; VA, AQ-R Verbal Aggression; ANG, AQ-R Anger; HO, AQ-R Hostility.

of the test. With regard to the temporal stability, our results are quite appropriate taking into account that our test–retest interval was of 4 months. Other studies have reported higher test–retest reliabilities, but with shorter time intervals [15]. So, these results suggest remarkable stability over the time.

When we assessed the differences between men and women, we expected to find differences between them. Men were found to be more aggressive in Physical Aggression than women, like in the original study of Buss and Perry [11]. In contrast, women were found to be more aggressive than men with regard to Anger and Hostility, whereas in the original study men showed themselves as more aggressive in all subscales except Anger. These results suggest that the expression of cognitive and emotional components of aggressiveness might be different in the Spanish context when compared to the American context. As García-León et al. [15] suggest this could be an effect attributable to the different role of social pressure upon men and women in Spain. Thus, women would express a broad range of emotions, rather than men, who would inhibit most of them. This pattern of results is also consistent with that reported by Andreu-Rodríguez et al. [1].

On the other hand, when we explored the differences between populations, we found that police students showed themselves as systematically less aggressive than the unselected

sample or offenders. It must be noticed that these students must pass an entrance examination, in which they are selected by presenting a low level of aggressiveness, among many other variables. With regard to the differences between offenders and the unselected sample volunteers, offenders showed significantly more aggressive in Physical Aggression and Hostility, which is consistent with the fact that inmates are more aggressive than the unselected sample. Actually, it is known that anger and hostility are related to impulsive aggression, and they have revealed as risk factors if violence in corrections [37].

Thus, the AQ-R seems to be a good tool to assess aggressiveness, and consequently, it can be useful in the violence risk assessment and prediction of violent behaviours in the future, at least in forensic contexts, where it shows significant correlations with the HCR-20 and PCL-R [13]. However, in order to check whether it is really a good predictor or not, further research is needed on specificity and sensitivity of this tool, using ROC curves and external criteria, such as the presence or absence of actual aggression.

### 5. Conclusion

To sum up, the results of the present study further support the use of the AQ-R, both in clinical and forensic contexts, and

when screening unselected groups of population. The scale's brevity is its main strength. So, it makes our instrument a fast and reliable tool in order to assess aggressiveness, as long as it can be utilised in a broad variety of contexts, where long time spending tests fail, due to time or to subject's abilities.

### Acknowledgements

This work was partially supported by grants BSO2003-08507 (Albert Maydeu-Olivares, PI) and SEC2001-3821-C05-

01 (Antonio Andrés-Pueyo, PI) both from the Spanish Ministry of Science and Technology. We acknowledge the School of Police of Catalonia and the Department of Justice of the Catalan Government for their support in the gathering of the data. We also gratefully acknowledge four anonymous reviewers for their valuable comments improving the article.

### Appendix 1

These items compose the Spanish version of AQ-R. All items are direct.

A mi me pasa que ...	Nunca	Casi nunca	A veces	Casi Siempre	Siempre
1 ... me doy cuenta que estoy en desacuerdo con los demás.	1	2	3	4	5
2 ... siento que la vida me ha tratado mal.	1	2	3	4	5
3 ... he amenazado a personas que conozco.	1	2	3	4	5
4 ... me pregunto porque me siento tan amargado/a.	1	2	3	4	5
5 ... tengo problemas para controlar mi mal genio.	1	2	3	4	5
6 ... mis amigos/as dicen que soy discudidor/ra.	1	2	3	4	5
7 ... me enfado rápidamente aunque se me pasa deprisa.	1	2	3	4	5
8 ... si me provocan mucho puedo llegar a pegar a cualquiera.	1	2	3	4	5
9 ... no puedo evitar discutir con los que no están de acuerdo conmigo.	1	2	3	4	5
10 ... creo que siempre son los otros los que consiguen las mejores oportunidades.	1	2	3	4	5
11 ... hay personas que me hacen enfadar tanto que llegaríamos a las manos.	1	2	3	4	5
12 ... me enfado mucho sin ninguna razón aparente.	1	2	3	4	5

### References

- [1] Andreu-Rodríguez JM, Peña-Fernández ME, Graña-Gómez JL. Adaptación psicométrica de la versión española del Cuestionario de Agresión. *Psicothema* 2002;14(2):476–82.
- [2] Barratt ES. Impulsiveness subtraits. Arousal and information processing. In: Spence JT, Itard CE, editors. *Motivation, emotion and personality*. North-Holland: Elsevier Science Publishers. 1985. p. 113–32.
- [3] Barratt ES, Slaughter L. Defining, measuring, and predicting impulsive aggression: a heuristic model. *Behav Sci Law* 1998;16:285–302 (1998).
- [4] Berry JW. Introduction to methodology. In: Triandis H, Berry JW, editors. *Handbook of cross-cultural psychology*. Boston: Allyn & Bacon; 1980. p. 1–28.
- [5] Book AS, Quinsey VL. Psychopaths: cheaters or warrior-hawks? *Pers Individ Dif* 2004;36:33–45.
- [6] Brislin RW. Translation and content analysis of oral and written materials. In: Triandis H, Berry JW, editors. *Handbook of cross-cultural psychology*. Boston: Allyn & Bacon; 1980. p. 389–444.
- [7] Browne MW. FITMOD: a computer program for calculating point and interval estimates of fit measures [computer program]. 1992.
- [8] Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, editors. *Testing structural equation models*. Newbury Park, CA: Sage; 1993. p. 132–62.
- [9] Bryant FB, Smith BD. Refining the architecture of aggression: a measurement model for the Buss–Perry Aggression Questionnaire. *J Res Pers* 2001;35:138–67.
- [10] Buss AH, Durkee A. An inventory for assessing different kinds of hostility. *J Consult Psychol* 1957;21:343–9.
- [11] Buss AH, Perry M. The Aggression Questionnaire. *J Pers Soc Psychol* 1992;63:452–9.
- [12] Buss AH, Plomin R. *A temperament theory of personality development*. New York: Wiley-Interscience; 1975.
- [13] Claix A, Pham TH. Evaluation of the HCR-20 violence risk assessment scheme in a Belgian forensic population. *Encephale* 2003;2004:447–53.
- [14] Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
- [15] García-León A, Reyes GA, Vila J, Pérez N, Robles H, Ramos MM. The Aggression Questionnaire: a validation study in student samples. *Span J Psychol* 2002;5(1):45–53.
- [16] Giancola PR. Alcohol-related aggression in men and women: the influence of dispositional aggressivity. *J Stud Alcohol* 2002;63(6):696–708.
- [17] Goveas JS, Csernasky JG, Coccaro EF. Platelet serotonin content correlates inversely with life history of aggression in personality-disordered subjects. *Psychiatry Res* 2004;126:23–32.
- [18] Kendell R, Jablenski A. Distinguishing between the validity and utility of psychiatric diagnosis. *Am J Psychiatry* 2003;160:4–12.
- [19] Kramp U. [Efecto del número de opciones de respuesta sobre las propiedades psicométricas de los cuestionarios de personalidad] Effects of the number of response options on the personality rating scales [dissertation]. University of Barcelona, Barcelona: Dept. of Personality; 2006.
- [20] Likert R. A technique for the measurement of attitudes. *Arch Psychol (Frankf)* 1932;140:5–55.
- [21] Likert R, Roslow S, Murphy GA. Simplified and reliable method of scoring the Thurstone attitude scales. *J Soc Psychol* 1934;5:228–38.
- [22] Luengo MA, Carrillo-de-la-Peña MT, Otero JM. The components of impulsiveness: a comparison of the I7 Impulsiveness Questionnaire and the Barratt Impulsiveness Scale. *Pers Individ Dif* 1991;12:657–67.
- [23] Miotto P, De Coppi M, Frezza M, Petretto D, Masala C, Pretti A. Suicidal ideation and aggressiveness in school-aged youths. *Psychiatry Res* 2003;120(3):247–55.
- [24] Miotto P, De Coppi M, Frezza M, Petretto D, Masala C, Pretti A. Eating disorders and aggressiveness among adolescents. *Acta Psychiatr Scand* 2003;108(3):183–9.

- [25] Morales-Vives F, Codorniu-Raga MJ, Vigil-Colet A. Características psicométricas de las versiones reducidas del cuestionario de agresividad de Buss y Perry. *Psicothema* 2005;17(1):96–100.
- [26] Muthén L, Muthén B. *MPLUS user's guide*. Los Angeles: Muthén and Muthén; 1998.
- [27] Patkar AA, Murray HW, Mannelli P, Gottheil E, Weinstein SP, Vergare MJ. Pre-treatment measures of impulsivity, aggression and sensation seeking are associated with treatment outcome for African-American cocaine-dependent patients. *Journal of J Addict Dis* 2004;23(2):109–22.
- [28] Pham TH. Psychometric assessment of the Buss and Perry Aggression Questionnaire (1992) and its relationship to the PCL-R in a Belgian prison population. *Cahiers de recherche*; 2004.
- [29] Seroczynski AD, Bergeman CS, Coccaro EF. Etiology of the impulsivity/aggression relationship: genes or environment? *Psychiatry Res* 1999;86: 51–7 (1999).
- [30] Si SX, Cullen JB. Response categories and potential cultural bias: effects of an explicit middle point in cross-cultural surveys. *Int J Organ Anal* 1998;6(3):218–30.
- [31] Sjöström N, Eder DN, Malm U, Beskow J. Violence and its prediction at a psychiatric hospital. *Eur Psychiatry* 2001;16:459–65.
- [32] Steiger JH. Structural model evaluation and modification: an interval estimation approach. *Multivariate Behav Res* 1990;25:173–80.
- [33] Swaminathan H, Rogers HJ. Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement* 1990;27:361–70.
- [34] Tremblay PF, Ewart LA. The Buss and Perry Aggression Questionnaire and its relation to values, the Big Five, provoking hypothetical situations, alcohol consumption patterns and alcohol expectancies. *Pers Individ Dif* 2005;38:337–46.
- [35] Tzeng OCS. A comparative evaluation of four response formats in personality ratings. *Educ Psychol Meas* 1983;43:935–50.
- [36] Vigil-Colet A, Lorenzo-Seva U, Codorniu-Raga MJ, Morales F. Factor structure of the Aggression Questionnaire among different samples and languages. *Aggress Behav* 2005;31:601–8.
- [37] Wang EW, Diamond PM. Empirically identifying factors related to violence risk in corrections. *Behav Sci Law* 1999;17:277–389.