Psychometric properties of the Automatic Thoughts Questionnaire–8 in Spain

Francisco J. Ruiz [Corresp., 1] , **Miguel A. Segura-Vargas** [1] , **Paula Odriozola-González** [2] , **Juan C. Suárez-Falcón** [3]

[1] Fundación Universitaria Konrad Lorenz, Bogotá, Colombia

[2] Universidad de Valladolid, Valladolid, Spain

[3] Universidad Nacional de Educación a Distancia, Madrid, Spain

Corresponding Author: Francisco J. Ruiz
Email address: franciscoj.ruizji@gmail.com

Background. The ATQ is a widely used instrument consisting of 30 items that measures the frequency of negative automatic thoughts. However, the extensive length of the ATQ could compromise its measurement efficiency in survey research. Accordingly, an 8-item shortened version of the ATQ has been developed. This study aims to analyze the validity of the ATQ-8 in Spanish samples. **Method.** The ATQ-8 was administered to a total sample of 1148 participants (302 undergraduates and 846 general online population). **Results.** The internal consistency across samples was good (alpha = .89). The one-factor model showed a good fit to the data (RMSEA = 0.089, 90% CI [0.089, 0.011], CFI = 0.98, NNFI = 0.97, and SRMR = 0.0483). The ATQ-8 scores were significantly associated with dysfunctional schemas, emotional symptoms, satisfaction with life, experiential avoidance, cognitive fusion, and generalized pliance. In conclusion, the Spanish version of the ATQ-8 showed good psychometric properties in Spain.

1

Psychometric properties of the Automatic Thoughts Questionnaire – 8 in two Spanish nonclinical samples

4

5

6 Francisco J. Ruiz[1], Miguel A. Segura-Vargas[1], Paula Odriozola-González[2], Juan C. Suárez-
7 Falcón[3]

8

9 [1] Faculty of Psychology, Fundación Universitaria Konrad Lorenz, Bogotá, Colombia
10 [2] Department of Psychology, Universidad de Valladolid, Valladolid, Spain
11 [3] Faculty of Psychology, Universidad Nacional de Educación a Distancia (UNED), Madrid,
12 Spain

13

14 Corresponding Author:
15 Francisco J. Ruiz[1]
16 Carrera 9 bis, N° 62-43, Bogotá, 110231, Colombia
17 Email address: franciscoj.ruizj@konradlorenz.edu.co

18

19

20

21

22

23

24

25

26

Abstract

27

28   **Background.** The ATQ is a widely used instrument consisting of 30 items that assess the

29   frequency of negative automatic thoughts. However, the extensive length of the ATQ could

30   compromise its measurement efficiency in survey research. Consequently, an 8-item shortened

31   version of the ATQ has been developed. This study aims to analyze the validity of the ATQ-8 in

32   Spanish samples.

33   **Method.** The ATQ-8 was administered to a total sample of 1148 participants (302

34   undergraduates and 846 general online population).

35   **Results.** The internal consistency across samples was adequate (alpha = .89). The one-factor

36   model demonstrated a good fit to the data (RMSEA = 0.10, 90% CI [0.089, 0.011], CFI = 0.98,

37   NNFI = 0.97, and SRMR = 0.048). The ATQ-8 scores were significantly associated with

38   emotional symptoms, satisfaction with life, dysfunctional schemas, cognitive fusion, experiential

39   avoidance, and generalized pliance. In conclusion, the Spanish version of the ATQ-8

40   demonstrated adequate psychometric properties in Spanish samples.

41

Introduction

42

43   Unipolar depression is characterized by sadness, irritability, or anhedonia, as well as a

44   loss of appetite, difficulty to sleep, fatigue, slowing of speech and action, and suicidal thoughts,

45   among others (American Psychiatric Association, 2013). The cognitive model proposed by Beck,

46   Rush, Shaw, and Emery (1979) states that the cognitive triad, integrated by a pattern of negative

47   thinking about the world, the future, and the self, is one of the pillars of depression. Within this

48   cognitive pattern, negative automatic thoughts play a crucial role and are defined as negative

49   self-statements (Beck et al., 1979).

50   The Automatic Thoughts Questionnaire is one of the most extensively used instruments

51   to measure negative automatic thoughts (ATQ; Hollon & Kendall, 1980). The ATQ is an

52   instrument consisting of 30 items with a 5-point Likert scale that assesses the frequency of

53   negative automatic thoughts experienced during the past week. Hollon and Kendall (1980) asked

54   312 undergraduates to recall dysphoric experiences and to report associated cognitions.

55   Afterward, the authors chose 100 representative cognitions and administered them to a second

56   sample. Through a cross-validation analysis, the authors retained 30 of the 100 original items.

57   These items significantly discriminated between clinical and nonclinical samples (Hollon &

58   Kendall, 1980).

59   Several studies have confirmed the temporal consistency, convergent and discriminant

60   validity, and excellent internal consistency of the ATQ (e.g., Chioqueta & Stiles, 2004; Hollon &

61   Kendall, 1980; Hollon, Kendall, & Lumry, 1986; Kazdin, 1990). The results of exploratory

62   factor analyses across different studies yielded factor structures with more than one factor (e.g.,

63   Bryant & Baxter, 1997; Cano-García & Rodríguez-Franco, 2002; Chioqueta & Stiles, 2006;

64   Deardorff, Hopkins, & Finch Jr, 1984; Ghassemzadeh, Mojtabai, Karamghadiri, &

65 Ebrahimkhani, 2005; Joseph, 1994; Kazdin, 1990; Oei & Mukhtar, 2008; Sahin & Sahin, 1992;

66 see reviews in Netemeyer et al., 2002; Zettle, Webster, Gird, Wagener, & Burdsal, 2013). Most

67 studies have obtained different factor solutions from the four factors shown by Hollon and

68 Kendall (1980). Netemeyer et al. (2002) mentioned that all studies found that a large proportion

69 of the variance was accounted for the first factor. Consequently, the results suggest that one

70 factor could underlie the 30 items of the ATQ. Moreover, most studies have only used the overall

71 score of the ATQ, which treats the scale as if it were only represented by one factor.

72   The extensive length of the ATQ could compromise its measurement efficiency in survey

73 research. Accordingly, Netemeyer et al. (2002) gathered two samples ($N = 434$ and $N = 419$) to

74 derive the 15- and 8-item reduced versions of the ATQ. Both versions of the questionnaire had a

75 single factor, with alphas of .96 and .92, respectively. Two additional cross-validation samples

76 ($N = 163$ and $N = 91$) also showed support for the 15-and 8-item reduced versions, which

77 suggests that the shortened versions of the ATQ are suitable alternatives to measure automatic

78 cognitions associated with depression (Netemeyer et al., 2002).

79   Following the study by Netemeyer et al. (2002), Ruiz, Suárez-Falcón, and Riaño-

80 Hernández (2017) analyzed the psychometric properties of the Spanish version of the ATQ-8 in a

81 Colombian sample of 1587 participants, including general population, a clinical sample, and

82 undergraduates. The analysis displayed good internal consistency across samples (alpha of. 89),

83 and the one-factor model obtained an adequate fit to the data (RMSEA = 0.083, 90% CI [0.074,

84 0.092]; CFI = .96; NNFI = .95). Additional factor analyses confirmed measurement invariance

85 across gender and samples (i.e., clinical and nonclinical samples). Furthermore, the mean scores

86 of the clinical sample were significantly higher than the scores of their nonclinical counterpart.

87   The results presented in Netemeyer et al. (2002) and Ruiz et al. (2017) indicate that the

88   ATQ-8 might be an excellent alternative to the original ATQ scale. However, the factor structure

89   and psychometric properties of the ATQ-8 have been analyzed only in two countries.

90   Accordingly, the current study aims to analyze the validity of the ATQ-8 in Spaniard samples.

91   This study is relevant because the original ATQ was only preliminarily validated in Spain by

92   Cano-García and Rodríguez-Franco (2002) in a sample of 205 individuals suffering from chronic

93   pain. Thus, there is scarce empirical evidence of the psychometric properties of the ATQ in

94   nonclinical samples in Spain.

95   This study analyzes the factor structure and psychometric properties of the ATQ-8 in two

96   nonclinical Spanish samples. The first sample consisted of 302 undergraduates and the second

97   one of 846 individuals from the general population.

98

99   **Materials & Methods**

100   The procedures followed in the research reported in the manuscript were approved by the

101   Bioethics Committee of Fundación Universitaria Konrad Lorenz (2016-021B). Written informed

102   consent was obtained from all participants in this study.

103   **Participants**

104   *Sample 1*. This sample consisted of 302 undergraduates (age range 18-61**,** $M = 26.18$, *SD*

105     $= 9.75$, 64.6% of females**)** from a Spanish university. Of the overall sample, 4.3% of the

106   participants were currently in treatment, 19.4% had received psychological or psychiatric

107   treatment, and 3.7% were taking psychotropic medication.

108   *Sample 2*. This sample consisted of 846 participants from general population, who

109   completed the instruments online (age range 18-72, $M = 35.14$, $SD = 11.39$, 75.7% of females).

110   Of the overall sample, 3.4% of participants had completed primary studies, 31% secondary

111   studies, and 55.6% were university graduates. Also, 12.8% of participants were currently in

112   treatment, 44.6% had received psychological or psychiatric treatment, and 12.9% were taking

113   psychotropic medication.

114   **Instruments**

115   *Automatic Thoughts Questionnaire – 8* (ATQ-8; Netemeyer et al., 2002; Spanish version

116   by Cano-García & Rodríguez-Franco, 2002). The ATQ-8 is the reduced version of the ATQ.

117   Through a Likert-type scale (5 = *all the time;* 1 = *not at all)*, it measures the frequency of

118   negative thoughts during the past week. Examples of items are "I'm so disappointed in myself,"

119   "I feel so helpless," "My future is bleak," and "I can't finish anything."

120   *Dysfunctional Attitude Scale-Revised* (DAS-R; de Graaf, Roelofs, & Huibers, 2009;

121   Spanish version by Ruiz et al., 2015). The DAS is a traditional instrument that measures

122   dysfunctional schemas. Its revised version (i.e., DAS-R) has 17 items, which are responded on a

123   7-point Likert-type scale (7 = *fully agree*; 1 = *fully disagree*), organized into two factors:

124   Performance evaluation and Dependency/ Perfectionism. Examples of the items are: "If a person

125   asks for help, it is a sign of weakness," "My happiness depends more on other people than it does

126   on me," "If I fail at my work, then I am a failure as a person," and "If others dislike you, you

127   cannot be happy." The DAS-R has shown a factor structure with two correlated factors and a

128   second-order factor and has also demonstrated adequate psychometric properties in Spanish and

129   Colombian samples (Ruiz, Suárez-Falcón, Barón-Rincón et al., 2016; Ruiz et al., 2015). In this

130   study, the DAS-R presented a Cronbach's alpha of .88 in Sample 1. According to the cognitive

131   model of depression, medium to strong correlations were expected between the DAS-R and the

132   ATQ-8.

133   *Depression, Anxiety, and Stress Scales – 21* (DASS-21; Lovibond & Lovibond, 1995;

134   Spanish version by Daza, Novy, Stanley, & Averill, 2002). The DASS-21 measures negative

135   emotional states experienced during the last week through 21 items on a 4-point Likert-type scale

136   (3 = *applied to me very much or most of the time*; 0 = *did not apply to me at all*). Examples of the

137   items are: "I couldn't experience positive feeling," "I felt close to panic," and "I found it difficult

138   to relax." The DASS-21 has shown a hierarchical factor structure consisting of three first-order

139   factors (Depression, Anxiety, and Stress) and a second-order factor. The latter can be considered

140   as an overall indicator of emotional symptoms (Ruiz, García-Martín, Suárez-Falcón, &

141   Odriozola-González, 2017). The DASS-21 has also presented good convergent and discriminant

142   validity and internal consistency. Alpha values in this study for the DASS-Total were .92 and .95

143   for Sample 1 and 2, respectively. The DASS-21 was administered because, in previous studies,

144   emotional symptoms and not only depression have been strongly associated with the frequency

145   of negative thoughts. Consequently, strong correlations were expected between the DASS-21

146   subscales and the ATQ-8.

147   *Satisfaction with Life Scale* (SWLS; Diener, Emmons, Larsen, & Griffin, 1985; Spanish

148   version by Atienza, Pons, Balaguer, & García-Merita, 2000). The SWLS measures self-perceived

149   well-being through 5 items, graded with a 7-point Likert-type scale (7 = *strongly agree*; 1 =

150   *strongly disagree*). Examples of items are "If I could live my life over, I would change almost

151   nothing," "In most ways, my life is close to my ideal," and "The conditions in my life are

152   excellent." The SWLS has demonstrated adequate convergent validity and psychometric

153   properties. Alpha values in the study were .84 and .89 for Samples 1 and 2, respectively.

154   Previous research has demonstrated that the frequency of negative thoughts is negatively

155   associated with life satisfaction (Ruiz et al., 2017). Medium to strong negative correlations were

156   expected between the SWLS and the ATQ-8.

157   *Acceptance and Action Questionnaire – II* (AAQ-II; Bond et al., 2011; Spanish version by

158   Ruiz, Langer, Luciano, Cangas, & Beltrán, 2013). The AAQ-II measures general experiential

159   avoidance through 7 items and a 7-point Likert-type scale (7 = *always*; 1 = *never true*). The

160   items evaluate the reluctance to experience unwanted emotions and thoughts as well as the

161   inability to be in the present moment and behave towards value-directed actions when

162   experiencing psychological discomfort. Examples of items are: "Emotions cause problems in my

163   life," "I worry about not being able to control my worries and feelings," and "It seems like most

164   people are handling their lives better than I am." The Spanish version by Ruiz et al. (2013)

165   demonstrated a one-factor structure and good psychometric properties in Spanish samples with

166   an overall alpha of .88. Alpha values in this study were .91 for both Sample 1 and Sample 2. The

167   AAQ-II was administered because prior research has obtained strong positive correlations

168   between ATQ scores and the AAQ-II (e.g., Ruiz & Odriozola-González, 2016).

169   *Cognitive Fusion Questionnaire* (CFQ; Gillanders et al., 2014; Spanish version by Ruiz,

170   Suárez-Falcón, Riaño-Hernández, & Gillanders, 2017). The CFQ measures cognitive fusion as

171   averaged across contexts through 7 items and a 7-point Likert-type scale (7 = *always*; 1 = *never*

172   *true*), where higher scores indicate a higher degree of cognitive fusion. Examples of the items

173   are: "I over-analyze situations to the point where it's unhelpful to me," "I get upset with myself

174   for having certain thoughts," and "I struggle with my thoughts." The English validation of the

175   CFQ has demonstrated to have good reliability, temporal stability, sensitivity to treatment

176   effects, a one-factor structure, and convergent, divergent, and discriminant validity. The Spanish

177   translation has proven to have similar psychometric properties (alpha = .92) and factor structure

178   to the original version (Ruiz, Suárez-Falcón, Riaño-Hernández, & Gillanders, 2017). In this

179   study, the CFQ obtained alphas of .90 and .93 for Samples 1 and 2, respectively. Medium to

180   strong positive correlations between the CFQ and the ATQ-8 were expected.

181   *Generalized Pliance Questionnaire* (GPQ; Ruiz, Suárez-Falcón, Barbero-Rubio, & Flórez,

182   2019). The GPQ is a questionnaire consisting of 18 items, graded on a 7-point Likert-type scale

183   (7 = *always true*; 1 = *never true*) that measures generalized pliance, defined as a pattern of rule-

184   governed behavior in which the individual's primary source of reinforcement is social whim.

185   Examples of the items are: "I care a lot about what my friends think of me," "My main goal in

186   life is to be recognized and respected by those around me," and "My decisions are very much

187   influenced by other people's opinions." In this study, the GPQ obtained an alpha of .92 and .95 in

188   Samples 1 and 2, respectively. Medium to strong positive correlations were expected between

189   the GPQ and the ATQ-8.

190   **Procedure**

191   For Sample 1, the instruments package was administered in the classrooms during a

192   regular class. In Sample 2, participants answered an online survey that was advertised through

193   social media (e.g., Facebook, institutional webpages, etc.). In both samples, participants provided

194   written informed consent. Participants in Sample 1 responded to the following instruments:

195   ATQ-8, DAS-R, DASS-21, SWLS, AAQ-II, CFQ, and GPQ. Participants in Sample 2 responded

196   to the same questionnaires except for the DAS-R. Once the participants completed the study, the

197   aims of the study were debriefed, and they were also thanked for their participation. No

198   incentives were provided to the participants.

199   **Statistical and Psychometric Analysis**

200   Before conducting factor analyses, the data from both samples were examined to find

201   missing values. However, no missing data were found. Corrected item-total correlations were

202   computed on SPSS 25$^{©}$ to find items that should be removed due to a low discrimination item

203   index (i.e., values below .30). McDonald's omega and Cronbach's alpha were conducted to

204   explore the ATQ-8 internal consistency with total sample (N = 1148) and providing percentile

205   bootstrap confidence intervals (CI) (Viladrich, Angulo-Brunet, & Doval, 2017). The MBESS

206   package in R was used to compute these coefficients (Kelley & Lai, 2012; Kelley &

207   Pornprasertmanit, 2016).

208   Because the ATQ-8 is responded on a 5-point Likert-type scale, an estimation method

209   appropriate for ordinal data was selected to conduct the CFA. Accordingly, a robust diagonally

210   weighted least square estimation method (Robust DWLS) was adopted using polychoric

211   correlations. These analyses were conducted with LISREL © (version 8.71, Jöreskog & Sörbom,

212   1999). For the one-factor model, the chi-square test and the following goodness of fit indexes

213   were calculated: (a) the root mean square error of approximation (RMSEA), (b) the comparative

214   fit index (CFI), (c) the non-normed fit index (NNFI), and (d) the standardized root mean squared

215   residual (SRMR). SRMR values below 0.05 reflect a very good fit to the data and values of 0.08

216   reflect a good fit to the data (Hu & Bentler, 1999; Kelloway, 1998). Kelloway (1998) suggested

217   that values of RMSEA of 0.10 represent an acceptable or modest fit, whereas Hu and Bentler

218   reduced the value to 0.08. Nevertheless, both guidelines suggest that a value of 0.05 reflects a

219   very good fit to the data. Regarding the CFI and NNFI, values above .95 show a good fit to the

220   data and above .90 indicate adequate-fitting models.

221   Following Jöreskog (2005) and Millsap and Yun-Tein (2004), additional CFAs were

222   conducted to assess for metric and scalar invariances across samples, gender, groupage (younger

223   or equal to 35 years and older than 35 years), and education level (primary and secondary studies

224   vs. university studies). Metric invariance means that item factor loadings are invariant across

225   samples, gender, groupage, and education level, whereas scalar invariance involves that item

226   intercepts are also invariant. Consequently, a comparison was conducted among the relative fits

227   of three increasingly restrictive models: the scalar invariance model, the metric invariance

228   model, and the multiple-group baseline model. In so doing, we compared the relative fit of three

229   increasingly restrictive nested models: the multiple-group baseline model (it allowed the

230   unstandardized factor loadings to vary across groups), the metric invariance model (it placed

231   equality of factor loadings across groups), and the scalar invariance model (it placed equality in

232   both the factor loadings and the item intercepts across groups). For the comparison model, the

233   indices of the CFI, NNFI, and RMSEA were compared among the nested models. Regarding the

234   selection of a model, the more constrained model was carefully chosen (i.e., second model versus

235   the first model, and third model versus the second model) if the following criteria proposed by

236   Cheung and Rensvold (2002) and Chen (2007) was fulfilled: (a) the difference in RMSEA

237   (ΔRMSEA) was lower than .01; (b) the differences in CFI (ΔCFI) and NNFI (ΔNNFI) were

238   higher or equal to -.01.

239   Descriptive data were also calculated. To explore gender differences in ATQ-8 scores, an

240   independent *t*-test was computed. Lastly, to evaluate convergent construct validity, Pearson

241   correlations between the ATQ-8 and the other instruments were calculated.

242

243   **Results**

244   **Descriptive data and psychometric quality of the items**

245   Table 1 displays the Spanish translation of the items of ATQ-8 with their corrected item-

246   total correlations for each sample and descriptive data. The eight items presented corrected item-

247   total correlation ranging from .55 to .74 for the overall sample and good discrimination indices.

248   INSERT TABLE 1 ABOUT HERE

249   Table 2 presents the alpha and omega coefficients of the ATQ-8 for Samples 1 and 2. The

250   alpha of the overall sample was .89 (95% CI [.88, .90]), whereas the omega was also .89 (95%

251   CI [.88, .90]). Table 2 also shows the descriptive data of the ATQ-8. There were no statistically

252   significant differences on the ATQ-8 scores between genders in Sample 1 (women: $M = 16.54$,

253    $SD = 6.92$; men: $M = 15.16$, $SD = 6.40$). However, in Sample 2, women showed higher scores on

254   the ATQ-8 than men (women: $M = 14.39$, $SD = 5.44$; men: $M = 14.58$, $SD = 5.35$).

255   INSERT TABLE 2 ABOUT HERE

256   **Validity evidence based on internal structure**

257   *Dimensionality*

258   The one-factor model obtained an adequate fit according to the goodness-of-fit indexes[1]:

259   χ^2 (20) $= 251.202$, $p < .01$; RMSEA $= 0.10$, 90% CI [0.089, 0.112], CFI $= 0.98$, NNFI $= 0.97$,

260   and SRMR $= 0.0483$. Figure 1 presents the results obtained from the completely standardized

261   solution of the one-factor model.

262   INSERT FIGURE 1 ABOUT HERE

263   *Measurement invariance*

264   Table 3 displays the results of the analysis of the scalar and metric invariance. Changes in

265   RMSEA, CFI, and NNFI were lower than .01 in all cases. Therefore, parameter invariance was

266   supported at both the scalar and metric levels across samples, gender, groupage, and education

267   level.

268 INSERT TABLE 3 ABOUT HERE

269 **Validity evidence based on relationships with other variables**

270 Table 4 shows that the ATQ-8 presented correlations with all of the other constructs that

271 were assessed in the expected direction: it presented positive correlations with dysfunctional

272 schemas (DAS-R), experiential avoidance (AAQ-II), emotional symptoms (DASS-21),

273 generalized pliance (GPQ), and cognitive fusion (CFQ); and negative correlations life

274 satisfaction (SWLS).

275 INSERT TABLE 4 ABOUT HERE

276
277 **Discussion**

278 While the ATQ has already been validated in Spain, to our best knowledge, no study has

279 analyzed the factor structure and psychometric properties of the ATQ-8 with Spanish samples.

280 This version has two main advantages over the original ATQ. Firstly, the factor structure of the

281 ATQ-8 is more simple and stable than the one of the original ATQ. Secondly, the ATQ-8 is

282 better suited to survey research and provides a considerably briefer assessment than the original

283 ATQ. Accordingly, this study aimed to explore the psychometric properties of the ATQ-8 in two

284 Spanish samples.

285 The analyses indicated that the Spanish version of the ATQ-8 showed good psychometric

286 properties in Spain. Concerning internal consistency, the ATQ-8 displayed an alpha of .89, and

287 the items had corrected item-total correlations ranging from .38 to .74. Confirmatory factor

288 analyses showed that the one-factor model presented a good fit to the data as in the previous

289 studies by Netemeyer et al. (2002) and Ruiz, Suárez-Falcón, et al. (2017). Also, the ATQ-8

290 showed metric and scalar measurement invariance across the type of sample (undergraduates and

291 general online population), gender, groupage (younger or equal than 35 years vs. older than 35

292   years), and education level (primary and secondary studies vs. university studies). These

293   analyses indicate that the ATQ-8 scores can be compared across these variables. Additionally,

294   the ATQ-8 demonstrated convergent validity, given the positive correlations found with

295   emotional symptoms, dysfunctional schemas, generalized pliance, experiential avoidance and

296   cognitive fusion, and the negative correlations with life satisfaction.

297   It is worth to mention some limitations of this study. Firstly, we did not collect data from

298   a clinical sample. This is a significant limitation because the ATQ was mainly designed to assess

299   clinical participants. Accordingly, further studies should analyze the psychometric properties of

300   the ATQ-8 in a clinical sample and, as in Ruiz, Suárez-Falcón, et al. (2017), to explore the

301   measurement invariance across clinical and nonclinical samples. Secondly, as this study did not

302   include a clinical sample, we were not able to analyze if the ATQ-8 can be used as a screening

303   measure to detect unipolar depression. Thirdly, the psychometric properties of the ATQ-8 were

304   analyzed in two convenience samples. Thus, the representativeness of the samples is uncertain.

305   Accordingly, further studies should be conducted with other Spaniard samples to confirm the

306   results of the current study. Fourthly, we did not explore the sensitivity to treatment. However,

307   note that the study by Ruiz, Suárez-Falcón, et al. (2017) showed that the ATQ-8 was sensitive to

308   treatment in a clinical study conducted in Colombia. Lastly, the percentage of women was

309   significantly higher than the percentage of men in the composition of the samples. However, the

310   finding of measurement invariance across gender reduces this limitation.

311   **Conclusions**

312   The findings of the current study are consistent with previous studies by Netemeyer et al.

313   (2002) and Ruiz, Suárez-Falcón, et al. (2017). Importantly, this study adds empirical evidence of

314   the adequate fit of the one-factor structure of the ATQ-8 and its measurement invariance across

315   gender, age, and education level. Further studies should try to replicate these findings in other

316   Spanish-speaking countries and analyze the measurement invariance of the ATQ-8 across

317   different cultures and countries.

318   In conclusion, this study showed that the ATQ-8 appears to be a reliable and valid

319   instrument in Spanish samples. Therefore, the ATQ-8 can be used in Spain as a less time-

320   consuming measure of negative automatic thoughts than the original ATQ.

321

322

323

324   **Acknowledgments**

325

326

327   **References**

328   American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders*

329   (DSM-5®). American Psychiatric Pub.

330   Atienza, F. L., Pons, D., Balaguer, I., & García-Merita, M. 2000. Propiedades psicométricas de

331   la Escala de Satisfacción con la vida en adolescentes [Psychometric properties of the

332   Satisfaction with Life Scale in adolescents]. *Psicothema, 12*, 314–319.

333   Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. 1979. *Cognitive therapy of depression.* New

334   York, NY: Guilford Press.

335   Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., … Zettle,

336   R. D. 2011. Preliminary psychometric properties of the Acceptance and Action

337   Questionnaire – II: A revised measure of psychological inflexibility and experiential

338   avoidance. *Behavior Therapy, 42*, 676–688. https://doi.org/10.1016/j.beth.2011.03.007

339   Bryant, F. B., & Baxter, W. J. 1997. The structure of positive and negative automatic cognition.

340   *Cognition & Emotion, 11*, 225-258. https://doi.org/10.1080/026999397379908

341   Cano-García, F. J., & Rodríguez-Franco, L. 2002. Evaluación del lenguaje interno ansiógeno y

342   depresógeno en la experiencia de dolor crónico [Assessment of anxious and depressive

343   self-talk in chronic pain experience]. *Apuntes de Psicología, 20*, 329–346.

344   Chen, F. F. 2007. Sensitivity of goodness of fit indexes to lack of measurement invariance.

345   *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 464–504.

346   https://doi.org/10.1080/10705510701301834

347   Cheung, G. W., & Rensvold, R. B. 2002. Evaluating goodness-of-fit indexes for testing

348   measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*,

349   233–255. https://doi.org/10.1207/S15328007SEM0902_5

350   Chioqueta, A. P., & Stiles, T. C. 2004. Norwegian version of the automatic thoughts

351   questionnaire: a reliability and validity study. *Cognitive Behaviour Therapy, 33*, 79–82.

352   https://doi.org/10.1080/16506070310016031

353   Chioqueta, A. P., & Stiles, T. C. 2006. Factor structure of the dysfunctional attitude scale (Form

354   A) and the automatic thoughts questionnaire: An exploratory study. *Psychological

355   Reports, 99*, 239–247. https://doi.org/10.2466/pr0.99.1.239-247

356   Daza, P., Novy, D. M., Stanley, M., & Averill, P. 2002. The Depression Anxiety Stress Scale-21:

357   Spanish translation and validation with a Hispanic sample. *Journal of Psychopathology

358   and Behavioral Assessment, 24*, 195–205. https://doi.org/10.1023/A:1016014818163

359   Deardorff, P. A., Hopkins, L. R., & Finch, A. J., Jr. 1984. Automatic Thoughts Questionnaire: A

360   reliability and validity study. *Psychological Reports, 55*, 708–710.

361   https://doi.org/10.2466/pr0.1984.55.3.708

362   de Graaf, L. E., Roelofs, J., & Huibers, M. J. 2009. Measuring Dysfunctional Attitudes in the

363   General Population: The Dysfunctional Attitude Scale (form A) Revised. *Cognitive*

364   *Therapy and Research, 33*, 345–355. https://doi.org/10.1007/s10608-009-9229-y

365   Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. 1985) The satisfaction with Life Scale.

366   *Journal of Personality Assessment, 49*, 71–75.

367   https://doi.org/10.1207/s15327752jpa4901_13

368   Ghassemzadeh, H., Mojtabai, R., Karamghadiri, N., & Ebrahimkhani, N. 2005. Psychometric

369   properties of a Persian-language version of the Automatic Thoughts Questionnaire: ATQ-

370   Persian. *International Journal of Social Psychiatry, 52*, 127–137.

371   https://doi.org/10.1002/da.20070

372   Gillanders, D. T., Bolderston, H., Bond, F. W., Dempster, M., Flaxman, P. E., Campbell, L., &

373   Remington, B. 2014. The development and initial validation of the Cognitive Fusion

374   Questionnaire. *Behavior Therapy, 45*, 83–101.

375   http://dx.doi.org/10.1016/j.beth.2013.09.001

376   Hollon, S. D., & Kendall, P. C. 1980. Cognitive self-statements in depression: Development of

377   an automatic thoughts questionnaire. *Cognitive Therapy and Research, 4*, 383–395.

378   https://doi.org/10.1007/BF01178214

379   Hollon, S. D., Kendall, P. C., & Lumry, A. 1986. Specificity of depressotypic cognitions in

380   clinical depression. *Journal of Abnormal Psychology, 95*, 52–59.

381   https://doi.org/10.1037/0021-843X.95.1.52

382   Hu, L. T., & Bentler, P. M. 1999. Cutoff criteria for fit indexes in covariance structure analysis:

383   Conventional criteria versus new alternatives. *Structural equation modeling: a*

384   *multidisciplinary journal, 6*, 1-55. https://doi.org/10.1080/10705519909540118

385   Jöreskog, K. G. 2005. *Structural equation modeling with ordinal variables using LISREL*.

386   Lincolnwood, IL: Technical report.

387   Jöreskog, K. G., & Sörbom, D. 1996. *PRELIS 2 user's reference guide: A program for*

388   *multivariate data screening and data summarization: A preprocessor for LISREL*.

389   Chicago, IL: Scientific Software International. Scientific Software International, Inc.

390   Jöreskog, K. G., & Sörbom, D. 1999. *LISREL 8.30*. Chicago, IL: Scientific Software

391   International.

392   Joseph, S. 1994. Subscales of the automatic thoughts questionnaire. *The Journal of Genetic*

393   *Psychology, 155*, 367–368. https://doi.org/10.1080/00221325.1994.9914786

394   Kazdin, A. E. 1990. Evaluation of the Automatic Thoughts Questionnaire: Negative cognitive

395   processes and depression among children. *Psychological Assessment: A Journal of*

396   *Consulting and Clinical Psychology, 2*, 73–79. https://doi.org/10.1037/1040-3590.2.1.73

397   Kline, R. B. 2005. *Principles and practice of structural equation modeling*. New York, NY:

398   Guilford Press.

399   Lovibond, P. F., & Lovibond, S. H. 1995. The structure of negative emotional states:

400   Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression

401   and Anxiety Inventories. *Behaviour Research and Therapy, 33*, 335–343.

402   https://doi.org/10.1016/0005-7967(94)00075-U

403   Millsap, R. E., & Yun-Tein, J. 2004. Assessing factorial invariance in ordered-categorical

404   measures. *Multivariate behavioral research, 39*, 479-515.

405   https://doi.org/10.1207/S15327906MBR3903_4

406   Netemeyer, R. G., Williamson, D. A., Burton, S., Biswas, D., Jindal, S., Landreth, S., …

407   Primeaux, S. 2002. Psychometric properties of shortened versions of the Automatic

408   Thoughts Questionnaire. *Educational and Psychological Measurement, 62,* 111–129.

409   https://doi.org/10.1177/0013164402062001008

410   Oei, T. P. S., & Mukhtar, F. 2008. Exploratory and confirmatory factor validation and

411   psychometric properties of the automatic thoughts questionnaire for Malays (ATQ-

412   Malay) in Malaysia. *Hong Kong Journal of Psychiatry, 18*, 92–100.

413   Ruiz, F. J., Langer, A. I., Luciano, C., Cangas, A. J., & Beltrán, I. 2013. Measuring experiential

414   avoidance and psychological inflexibility: The Spanish translation of the Acceptance and

415   Action Questionnaire. *Psicothema, 25*, 123–129.

416   https://doi.org/10.7334/psicothema2011.239

417   Ruiz, F. J., & Odriozola-González, P. 2016. The role of psychological inflexibility in Beck's

418   cognitive model of depression. *Anales de Psicología, 32*, 441–447.

419   https://doi.org/10.6018/analesps.32.2.214551

420   Ruiz, F. J., Suárez-Falcón, J. C., Barbero-Rubio, A., & Flórez, C. L. 2019. Development and

421   initial validation of the Generalized Pliance Questionnaire. *Journal of Contextual*

422   *Behavioral Science, 12*, 189–198. https://doi.org/10.1016/j.jcbs.2018.03.003

423   Ruiz, F. J., Suárez-Falcón, J. C., Barón-Rincón, D., Barrera-Acevedo, A., Martínez-Sánchez, A.,

424   & Pena, A. 2016. Factor structure and psychometric properties of the Dysfunctional

425   Attitude Scale Revised in Colombian undergraduates. *Revista Latinoamericana de*

426   *Psicología, 48*, 81–87. https://doi.org/10.1016/j.rlp.2015.10.002

427   Ruiz, F. J., Suárez-Falcón, J. C., Odriozola-González, P., Barbero-Rubio, A., López-López, J. C.,

428   Eisenbeck, N., ... & Gil, E. 2015. Factor structure and psychometric properties of the

429   Spanish version of the" Dysfunctional Attitude Scale-Revised". *Psicología Conductual,*

430   *23*, 287–303.

431   Ruiz, F. J., Suárez-Falcón, J. C., & Riaño-Hernández, D. 2017. Validity evidence of the Spanish

432   version of the automatic thoughts questionnaire–8 in Colombia. *The Spanish Journal of*

433   *Psychology, 20*, e12. https://doi.org/10.1017/sjp.2017.7

434   Ruiz, F. J., Suárez-Falcón, J. C., Riaño-Hernández, D., & Gillanders, D. 2017. Psychometric

435   properties of the Cognitive Fusion Questionnaire in Colombia. *Revista Latinoamericana*

436   *de Psicología, 49*, 80–87. http://dx.doi.org/10.1016/j.rlp.2016.09.006

437   Sahin, N. H., & Sahin, N. 1992. Reliability and validity of the Turkish version of the Automatic

438   Thoughts Questionnaire. *Journal of Clinical Psychology, 48*, 334–340.

439   https://doi.org/10.1002/1097-4679(199205)48:3<334::AID-JCLP2270480311>3.0.CO;2-

440   P

441   Zettle, R. D., Webster, B. K., Gird, S. R., Wagener, A. L., & Burdsal, C. A. 2013. Factor

442   structure of the Automatic Thoughts Questionnaire in a clinical sample. *International*

443   *Journal of Cognitive Therapy, 6*, 280–291. https://doi.org/10.1521/ijct.2013.6.3.280

Table 1(on next page)

Item Description, Corrected Item-Total Correlations, Descriptive Data

PeerJ

1   Table 1.

2   *Item Description, Corrected Item-Total Correlations, Descriptive Data*

Item number and description	Corrected item-total correlation		
	Sample 1 Undergraduates	Sample 2 General population online	Overall Sample
1. No soy Bueno [I'm no good].	.51	.58	.56
2. ¡Soy tan decepcionante hasta para mí mismo! [I'm so disappointed in myself].	.72	.74	.74
3. ¿Qué es lo que funciona mal en mí? [What's wrong with me?].	.67	.75	.72
4. Soy un inútil, no valgo para nada [I'm worthless].	.62	.71	.70
5. Me siento tan impotente, tan desamparado [I feel so helpless].	.53	.74	.70
6. Algo tiene que cambiar [Something has to change].	.58	.71	.67
7. Mi future es un desierto [My future is bleak].	.48	.70	.66
8. No consigo terminar nada de lo que empiezo [I can't finish anything].	.38	.59	.55

3

4

Table 2(on next page)

Coefficient Alpha and Omega, and Descriptive Data across Samples

1   Table 2.

2   *Coefficient Alpha and Omega, and Descriptive Data across Samples*

	Sample 1: Undergraduates ($N = 302$)	Sample 2: General population online ($N = 846$)	Overall Sample ($N = 1148$)
Alpha [95% CI]	.83 [.80, .85]	.90 [.89, .91]	.89 [.88, .90]
Omega [95% CI]	.83 [.78, .86]	.90 [.89, .91]	.89 [.88, .90]
Mean score (*SD*)	14.46 (5.40)	16.22 (6.80)	15.76 (6.50)

3

4

Table 3(on next page)

Table 3

Table 3. *Metric and Scalar Invariance across Sample, Gender, Group Age, and Education Level*

1 Table 3.

2 *Metric and Scalar Invariance across Sample, Gender, Group Age, and Education Level*

Model	RMSEA	ΔRMSEA	CFI	ΔCFI	NNFI	ΔNNFI
Measurement invariance across sample						
MG Baseline model	.0983		.982		.975	
Metric invariance	.1000	-.0017	.978	- .004	.974	- .001
Scalar invariance	.0984	.0016	.976	- .002	.975	.001
Measurement invariance across gender						
MG Baseline model	.101		.980		.973	
Metric invariance	.0915	.0095	.981	.001	.977	.004
Scalar invariance	.0903	.0120	.979	-.002	.978	.001
Measurement invariance across group age						
MG Baseline model	.1001		.979		.971	
Metric invariance	.1047	-.0046	.973	-.006	.968	-.003
Scalar invariance	.1089	-.0042	.966	-.007	.965	-.003
Measurement invariance across education level						
MG Baseline model	.1013		.981		.974	
Metric invariance	.1046	-.0033	.976	-.005	.972	-.002
Scalar invariance	.1033	.0013	.974	-.002	.973	.001

Table 4(on next page)

Pearson Correlations between the ATQ-8 Scores and Other Relevant Self-report Measures

1 Table 4.

2 *Pearson Correlations between the ATQ-8 Scores and Other Relevant Self-report Measures*

Measure	S	N	*r* with ATQ-8
DAS-R	1	302	.43*
DASS – Total	1	302	.60*
	2	846	.74*
DASS – Depression	1	302	.61*
	2	846	.78*
DASS – Anxiety	1	302	.47*
	2	846	.56*
DASS – Stress	1	302	.49*
	2	846	.62*
AAQ-II	1	302	.59*
	2	846	.70*
CFQ	1	302	.63*
	2	846	.65*
GPQ	1	302	.31*
	2	846	.48*
SWLS	2	846	-.63*

3 *Note*. AAQ-II: Acceptance and Action Questionnaire – II; ATQ-8: Automatic Thoughts Questionnaire – 8; CFQ =
4 Cognitive Fusion Questionnaire; DAS-R: Dysfunctional Attitude Scale – Revised; DASS: Depression, Anxiety, and
5 Stress Scales – 21; GPQ = Generalized Pliance Questionnaire; SWLS: Satisfaction with Life Scale. *p < .001.
6

7

Figure 1

Completely standardized solution of the ATQ-8 one-factor model.