Acquiescence and Dogmatism: Impact and relations across two different item formats

The literature suggests that (a) responses to typical Dogmatism items are non-negligibly impacted by acquiescent responding, and (b) the dimension of Dogmatism is related to Acquiescence conceptualized as an individual-differences variable. In this article we first rigorously assessed these two basic hypotheses by using a recently developed confirmatory factor-analytic model. Second, we assessed the extent to which the impact and relations were consistent across a binary format and a graded format. Finally, the convergent validity of the Acquiescence scores across the two formats was assessed. In both datasets the Dogmatism items were found to be impacted by Acquiescence, and the relation Acquiescence-Dogmatism was strong in the expected direction. Furthermore, the Acquiescence scores were very consistent across the two different formats. The theoretical implications of the results and future lines of research are discussed.

Keywords: Personality Measurement, Acquiescent Responding, Dogmatism, Item Response Format, Balanced Scales, Factor Analysis, Marker Variables.

1. Introduction

Acquiescence (ACQ), defined as the tendency to endorse or agree with an item regardless of its content, has been viewed in two opposite ways in personality theory: Firstly, as a nuisance variable that is mostly test dependent and which has no substantive interest (Ray, 1970); secondly, as an individual-differences variable (IDV) that can provide useful information about the respondent (Couch & Keniston, 1960, McGee, 1962). From this second viewpoint, ACQ might have the properties of a personality state or a trait. If the latter were to be the case, then ACQ measures would have a certain degree of internal consistency, stability over time, generalizability across situations and measurement instruments, and convergent validity with respect to other relevant measures (e.g. Ferrando, Condon, & Chico, 2004, McGee, 1962).

In recent decades a certain consensus seems to have been reached regarding the status of ACQ as a variable and its state/trait properties. Acquiescent responding (AR) is regarded as an item × person process in which certain item characteristics interact with an enduring disposition of the respondent which minimally fulfils the trait requirements (Condon, Ferrando, & Demestre, 2006, Danner, Aichholzer & Rammstedt, 2015). At the same time, however, ACQ scores usually show a high degree of specificity. So, the consensus is be that ACQ is an IDV that behaves like a 'weak' trait and which has both state-like and trait-like properties (Danner, Aichholzer & Rammstedt, 2015, Ferrando, Condon & Chico, 2004). This variable is also generally thought of as dimensional, and modeled as a bipolar factor with a positive pole of tendency to agree and a negative pole of tendency to dissent (DiStefano, Morgan & Motl, 2012).

The consensus above leaves many open questions of which this article focuses on two. The first question concerns the relation between ACQ and the content trait that a

personality test intends to measure. More in detail, if both content and ACQ are viewed as dimensions and modeled as common factors, then a basic question is whether these factors are correlated (i.e. oblique factors) or not (i.e. orthogonal factors). The second question refers to the impact and consistency of AR across different item response formats. If ACQ is viewed as an IDV, then, a certain amount of across-format consistency is to be expected. However, in the interactive process referred to above, some item characteristics, mainly the item format, are likely to modulate the effects of ACQ on the item responses. This second question appears to have been addressed only partially in the literature, and previous research has only focused on the impact of ACQ on graded-response scales as a function of labeling, number of categories, and the use or not of a midpoint option (Moors, Kieruj, & Vermunt, 2014, Weijters, Cabooter & Schillewaert, 2010).

Regarding the basic first question above, the factor-analytic approaches proposed to date have modelled ACQ and content as orthogonal factors (Ferrando, Lorenzo-Seva & Chico, 2003, Savalei & Falk, 2014), which is expected to be appropriate for many personality traits. (e.g. Messick & Frederiksen 1958, McCrae & Costa, 1983). However, specific traits such as impulsivity, conscientiousness, conformity, external locus of control, open-mindedness, and authoritarism have been theoretically related to ACQ in the personality literature (e.g. Couch & Keniston, 1960, DiStefano, Morgan & Motl, 2012, Gudjonsson & Young, 2011). If this is the case, then, estimating the direction and strength of the relations between ACQ and theoretically related traits is of clear interest (McGee, 1962). This estimation is not feasible with the existing orthogonal FA models mentioned above. However, Ferrando, Morales-Vives, and Lorenzo-Seva (2016) have recently developed a testable oblique FA model that will serve as a basis for the present study.

1.1 Acquiescence and Dogmatism: A Review of Impact and Relations

Rokeach developed his original Dogmatism scales as one-way measures in which all the items measured in the same direction of Dogmatism or closed-mindedness, and, in his opinion, were essentially unaffected by AR (e.g. Rokeach, 1967). The literature review, however, suggests that the impact of AR on Dogmatism scales is generally non-negligible (Lichtenstein, Quinn & Hover, 1961, McGee, 1962, Peabody, 1966). With regards to the ACQ-Content relation, the bipolar dimension open-mindedness vs. Dogmatism is thought to be related to ACQ in the direction that the more dogmatic the respondent is, the more prone he/she is to engage in AR (Couch & Keniston, 1960, McGee, 1962, Messick & Frederiksen, 1958).

The two summaries above are mostly based on purely empirical results obtained from separate ACQ scales (see Ferrando, Condon & Chico, 2004). However, some theoretical explanations have also been offered. As far as the impact is concerned, authors such as Peabody (1966) and Ray (1979) noted that the typical item stems in Dogmatism scales are long and relatively complex statements which, in some cases, can also be somewhat ambiguous. It is these that appear to be the item characteristics that are most prone to eliciting AR (Condon, Ferrando & Demestre, 2006). As for the ACQ-Dogmatism relation, Couch and Keniston (1960) believed that acquiescent respondents generally have high levels of diffuse anxiety and that they seek dependency on external figures (or rigid beliefs) for support. Messick and Frederiksen (1958), on the other hand, offered an alternative explanation: that acquiescent respondents tend to lack critical capacity so they stick to rigid beliefs. Overall, we note that the evidence for the two points raised above is not based on rigorous applications, and that the theoretical

explanations that support them are only tentative. However, both are submitted to be plausible, and will be taken as the working hypotheses in the present study.

The assumption that Dogmatism scales are prone to be impacted by AR has prompted some authors to develop balanced measures in which some of the items measure in the direction of Dogmatism or closed-mindedness and the others measure in the direction of flexibility or open-mindedness (Ray, 1970,1974,1979, Altemeyer, 1996). A scale of this type has been used in the present study.

1.2 Objectives and predictions

This research has two groups of objectives. The objectives in the first group are to assess (a) the impact of ACQ on a Dogmatism measure, and (b) the latent correlation between the dimensions of ACQ and Dogmatism. Given the review above, our prediction regarding the first objective is as follows. The prime determinant of the responses to the Dogmatism items will be the dimension that is to be measured. However, ACQ will have a non-negligible impact on these responses, so an identifiable dimension of ACQ will emerge as a secondary factor. As for as objective (b), we expect the dimension of Dogmatism to be related to the ACQ dimension in the direction that high Dogmatism levels will be associated to greater proneness to engage in AR.

The second group of objectives has two parts. The first part aims to assess the extent to which results (a) and (b) above generalize across two different item response formats. The second part assesses the consistency of the ACQ scores across the two different formats. The two types of format we consider are (a) the binary format and (b) the graded response format. While the graded format is possibly the most common in personality measurement, the binary format is still popular (Eysenck's scales are a good example) and has non-negligible advantages (e.g. McDonald, 1999). Because the binary

format is fully labelled and has no neutral response option, we have used a graded format that is equivalent in these two aspects: a 6-point fully-labelled response format.

We turn now to the predictions in this second group of objectives. If ACQ is mostly related to the direction component of the response (i.e. agree vs. disagree; see Peabody, 1962), then (a) the impact of AR would be expected to be about the same in both formats, and (b) a high degree of across-format consistency would be expected for this dimension. Hypothesis (a) was put forward by Weijters, Cabooter & Schillewaert (2010) and the empirical findings appear to support it (Moors, Kieruj & Vermunt, 2014, Weijters, Cabooter & Schillewaert, 2010). So, as a starting point we adopt the parsimonious direction-component hypothesis, and we expect the following results. First, the factor structures and inter-factor correlation will be essentially the same in both formats. Second, the ACQ scores will show a high degree of consistency across the two formats.

1.3 Conceptual description of the model

Ferrando, Lorenzo-Seva, and Chico (2003) proposed a factor-analytic (FA) model for controlling and assessing AR in which ACQ was modelled as a second common factor. The model is intended to be used with a balanced scale, and can be identified by assuming that content and ACQ are orthogonal.

The present research uses a recent oblique extension of the model above (Ferrando, Morales-Vives & Lorenzo-Seva, 2016). Identification of this new model requires additional information, which can be obtained from a set of factor-pure item markers (e.g. Comrey & Lee, 1992) that help identify the ACQ factor. Overall, the model can be considered as a restricted (confirmatory) bidimensional FA model, in which the loadings

of the markers on the content factor are set to zero, and which provides a directly interpretable solution.

For the model to function appropriately the subset of markers must be well chosen and the scale fully balanced. For the first requirement, suitable markers can be found in specific ACQ scales (Ferrando, Condon & Chico, 2004). As for the second, in a fully-balanced scale, half of the items measure in one direction of the content trait whereas the other half measure in the opposite direction. However, in order to avoid the problems associated with negatively-worded items (e.g. Barnette, 2000) ideally all the items should be directly or positively worded (Hofstee, ten Berge & Hendricks, 1998). In the application of the model, half of the item scores are reverse-scored before the analysis is carried out, so all the item scores go in the same direction as the trait. With this setting, if the model is correct, all the loadings on the first factor (identified as the content factor) should be positive, and generally larger than those on the second factor. The second factor should be perfectly bipolar, and the negative loadings must coincide with the items that are reverse-scored (because in these items the tendency to agree implies a lower transformed score).

The model is applied in two stages. In the first stage, the items are calibrated by fitting the FA solution described and assessing its appropriateness. In the second stage (scoring) the calibration results (factor loading and inter-factor correlation) are taken as fixed and known, and are used for scoring individuals. If the model is correct, these scores are ‘clean’ measures of the corresponding dimension. So, the content scores are free of ACQ, and the ACQ scores are pure measures of this dimension. Finally, the reliability for both content and ACQ scores is also estimated. These estimates are relevant for two reasons. First, a minimum amount of reliability is a pre-requisite for any further validity assessment (in our case across-format consistency). Second, in a

well-designed scale the reliability of the content scores should be substantially larger than that of the ACQ scores (e.g. Billiet & McClendon, 2000). Further technical details are provided below, and a full methodological discussion can be found in Ferrando, Morales-Vives and Lorenzo-Seva (2016).

2. Method

2.1 Participants and procedure

A total of 440 undergraduate students from the Psychology and Education faculties of a Spanish university took part voluntarily in the study. Of these, 394 responded the questionnaire in graded format, 396 responded the binary version, and 350 were present at both administrations. The mean age was 21.5 (s.d.=5.3) and 81% were female. The calibration of the graded data was based on the 394 respondents available in this condition. The calibration of the binary data was based on the 396 respondents available in this condition. Finally, the analyses based on the individual scores require the binary and graded scores to be correlated. So, they were based on the 350 respondents for whom data were available in both conditions.

The questionnaires were administered twice with an interval of four weeks between the first and the second administration, and the order of presentation (binary-graded, or graded-binary) was balanced. In all cases, administration was in paper-and-pencil mode, and the questionnaires were completed voluntarily in classroom groups of 25 to 60 students and administered by the same person. The administration was anonymous, and the respondents had to provide their gender, date of birth, and preferred colour (the last two particulars were used to pair the questionnaires in both formats).

2.2 Measures

We used a 22-item questionnaire made up of 18 content items and 4 ACQ markers which were interspersed between the content items. The content items were taken from a Spanish version of Ray's (1970, 1974, 1979) balanced Dogmatism scale (BDS), and were fully balanced, with nine items measuring in the direction of Dogmatism and nine in the direction of open-mindedness. The Dogmatism-oriented items were then reversed, so the direction of the scoring was that a higher score means a higher level of flexibility or open-mindedness. The markers were taken from Couch and Keniston's (1960) Agreement Response Scale and were selected among the "cleanest" ACQ indicators obtained in Hundleby's (1966) analysis. The graded BDS version used a 6-point fully-labeled response format with "totally disagree" and "totally agree" endpoints. The binary version used a "True" "False" format. The ordered abbreviated item stems are shown in table 1.

(Please insert table 1 here.)

2.3 Analyses

In the calibration stage the bidimensional oblique model was fitted separately to the binary and graded datasets and model appropriateness was assessed. The criterion of estimation was unweighted least squares (ULS). Appropriateness was assessed by using two measures of fit which are reasonable under ULS estimation: the root-mean-squared residual (RMSR) and the gamma-goodness of fit index (GFI; see McDonald, 1999).

Provided the model was appropriate, the factor loadings and inter-factor correlation estimated at the calibration stage were taken as fixed and known, and used to obtain (a) individual factor scores and (b) the reliability estimates corresponding to

these scores in the second stage. As proposed in Ferrando et al. (2016) the individual scores were Bayes EAP estimates, and the reliability estimates were the marginal reliabilities of the EAP scores (see Ferrando et al. 2016, for details). This basic analysis was carried out using a non-commercial program developed by the authors, and which is freely available from them.

Finally, in the validity (i.e. across-format) analysis, the raw and disattenuated correlations between the binary-based and the graded-based factor scores were obtained in order to assess the degree of across-format consistency.

3. Results

3.1 Calibration stage

In both, the binary and the graded datasets, the fit of the bi-dimensional oblique model was reasonably good. For the binary data, the RMSR and the GFI were respectively 0.06 and 0.94. For the graded-response data they were 0.05 and 0.95. The corresponding patterns are in the upper part of table 2.

(Insert table 2 here.)

Regarding the factor patterns, in both cases the loadings on the first factor are all positive, as expected for the Dogmatism factor. However, they are only moderate (average loadings are 0.34 in the binary pattern and 0.36 in the graded pattern). In both formats, items C3, C8 and C12 are poor measures of the dimension. The Burt-Tucker coefficient of factorial congruence between the binary and the graded solutions for this

factor is 0.96, which suggests that the same factor is identified in both formats (Lorenzo-Seva & Ten Berge, 2006). As for the second factor, in both cases it is reasonably well identified by the markers, and has the expected bipolar structure for the 18 Dogmatism items. In the graded case, the agreement of the negative signs with the reverse-scored items is perfect. In the binary case the exception is item 21 and this considerably lowers the congruence value, which is 0.83. In both solutions, the loadings of the Dogmatism items on the second factor are low, which suggests that ACQ, although it has a non-negligible impact, is a residual factor. For the 18 Dogmatism items the second-factor loadings are larger in absolute value in the graded format (average 0.24) than in the binary format (average 0.17).

The estimated inter-factor correlation in both formats is provided in Table 2, together with the confidence band (68% confidence interval). There is considerable agreement with both estimates: a substantial negative correlation, which agrees with the theoretical expectations (i.e., the less flexible, or more dogmatic the respondent, the more prone he/she is to engage in AR). This result supports the identification of the second factor as the ACQ factor in both formats.

3.2 Individual Scoring

Table 3 summarizes first the EAP-based reliability estimates corresponding to the content and ACQ factor scores in both versions. The estimated reliability of the content scores is about the same in both formats, and is at the low end of what can be expected with a typical 18-item personality test. As was also expected, the reliability of the ACQ scores is lower than that of the content scores even when it is based on 22 items. However, the ACQ estimates in table 3 are quite respectable for a ‘residual’ dimension such as ACQ.

Finally, the estimated ACQ reliability is larger in the graded-response case.

(Insert table 3 here.)

Overall, the fact that in the graded-response data (a) the structure of the second factor is clearer, (b) the absolute loadings are larger, and the (c) the reliability of the corresponding scores is also larger, suggests that the impact of ACQ is greater with this format. It cannot be discarded that this result might be due (at least in part) to the traditional problems that arise in some conditions when binary data is factor-analyzed (e.g. McDonald, 1999). However, we note that results (a), (b) and (c) above apply only to the second factor. The quality of the first factor is about the same in both formats.

Table 4 shows the product-moment correlations between the binary-based and graded-based EAP scores in each factor together with the Bootstrap-based confidence bands (68% confidence intervals). Given that they assess the consistency of the content and the ACQ measures across the two different response formats, these correlations can be interpreted as convergent validity coefficients (e.g. Ferrando, Condon, & Chico, 2004). The upper panel of the table shows the raw validity coefficients. The lower panel shows the disattenuated estimates corrected using the reliability estimates reported in table 3.

(Insert table 4 here.)

Regarding table 4, the disattenuated validity for the content scores is rather high, which agrees with previous results that, essentially, the same Dogmatism dimension is measured in both formats. As expected the validity coefficients regarding the ACQ factor are lower than the content coefficients, but they are quite respectable, which

suggests that the ACQ scores have a substantial degree of consistency across formats and also over time (remember that the time interval between both administrations was four weeks).

4. Discussion and Conclusions

This article has addressed a series of issues concerning ACQ as a personality dimension. Some of these issues have already been discussed in the literature (e.g. McGee, 1962) while others appear to be new. The novelty of the issues that are not new is that here they are rigorously assessed using a testable factor-analytic model.

The previously considered issues refer to the impact of ACQ on Dogmatism and the relations between both dimensions, and our results agree with expectations. First, Dogmatism items (at least of the type considered here) appear to be non-negligibly affected by AR. Second, the dimensions of Dogmatism and ACQ appear to be substantially related in the expected direction. The impact of AR can be assessed and controlled with the modelling procedures used in our study. However, this control requires a well-balanced Dogmatism scale. So, the development of balanced scales of this type seems a worthwhile future aim, and more so given that the amount of measurement error in the scale we used here is considerable. Our result could also serve as the basis for future theoretical research into the relation between ACQ and Dogmatism.

The most novel results have to do with the across-format relations. We found that (a) the same basic FA solution was essentially obtained in the binary and the graded-response cases; (b) the estimated ACQ-Dogmatism relation was very similar in both formats; and (c) the binary-based and the graded-based ACQ factor scores showed a noticeable degree of across-format consistency. These results have important

implications. Thus, result (b) provides further support to our working hypothesis that the tendency to engage in AR is related to higher levels of closed-mindedness. Result (c) complements previous findings (e.g. Ferrando, Condon, & Chico, 2004): not only does ACQ appear to have a certain consistency across different measures and over time, but also across different formats and over time. So, our new results provide further support to the conceptualization of ACQ as a (possibly weak) personality trait.

Result (a) above suggests that the direction component of the response (i.e. agree vs. disagree) is the main determinant of AR. However, we also found that the ACQ factor was stronger, better defined, and more reliable in the graded-response case. Whether this result reflects method effects or indicates that the intensity component of the response (see Peabody, 1962) plays a role in the AR process is an issue that deserves further study.

Our study has several limitations. To start with, too many generalizations cannot be made on the basis of a single study, and intensive research is needed to assess the validity and generalizability of our results. However, we would like to stress that, in general, the results seem to be strong and quite meaningful. Second, we used a sample of undergraduate students, and this is clearly a limitation for generalizing results. In this respect, however, two points must be considered. First, several authors (Peabody, 1966, Ray, 1970, 1974, Rokeach, 1967) have pointed out that university students seem to be the type of sample in which Dogmatism scales work best, and it seems reasonable to start with the type of sample that is most likely to provide clearer results. Second, our study uses a two-wave balanced design with a fixed retest interval, which is difficult to implement with samples obtained from the general population. In any case, a clear future aim is to extend the study to different types of sample.

References

Altemeyer, B. (1996). *The authoritarian specter*. Cambridge, MA: Harvard University Press.

Barnette, J.J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, *60*, 361-370.

Billiet, J.B., & McClendon, M.J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling*, 7, 608-628.

Comrey, A.L., & Lee, H.B. (1992). *A first course in factor analysis*. Psychology Press.

Condon, L., Ferrando, P.J., & Demestre, J. (2006). A note on some item characteristics related to acquiescent responding. *Personality and individual differences*, *40*, 403-407.

Couch, A., & Keniston, K. (1960). Yeasayers and naysayers: Agreeing response set as a personality variable. *Journal of Abnormal and Social Psychology*, *60*, 151-174.

Danner, D., Aichholzer, J., & Rammstedt, B. (2015). Acquiescence in personality questionnaires: Relevance, domain specificity, and stability. *Journal of Research in Personality*, *57*,119-130.

DiStefano, C., Morgan, G.B., & Motl, R.W. (2012). An examination of personality characteristics related to acquiescence. *Journal of Applied Measurement, 13*, 41-56.

Ferrando, P.J., Condon, L. & Chico, E. (2004). The convergent validity of acquiescence: an empirical study relating balanced scales and separate acquiescence scales. *Personality and Individual Differences*, *37*, 1331-1340.

Ferrando, P.J., Lorenzo, U. & Chico, E. (2003). Unrestricted factor analytic procedures for assessing acquiescent responding in balanced, theoretically unidimensional personality scales. *Multivariate Behavioral Research*, *38*, 353-374.

Ferrando, P.J., Morales-Vives, F., & Lorenzo-Seva, U. (2016). Assessing and controlling acquiescent responding when acquiescence and content are related: A comprehensive factor-analytic approach. . *Structural Equation Modeling*, (in press).

Gudjonsson, G.H., & Young, S. (2011). Personality and deception. Are suggestibility, compliance and acquiescence related to socially desirable responding?. *Personality and Individual Differences*, *50*, 192-195.

Hofstee, W.K.B. ten Berge, J.M.F. & Hendriks, A.A.J. (1998). How to score questionnaires. *Personality and Individual Differences*, *25*, 897-909.

Hundleby, J.D. (1966). The Construct Validity of a Scale of Acquiescence. *British Journal of Social and Clinical Psychology*, *54*, 290-298.

Lichtenstein, E., Quinn, R., & Hover, G.L. (1961). Dogmatism and acquiescent response set. *The Journal of Abnormal and Social Psychology*, *63*, 636.

Lorenzo-Seva, U., & Ten Berge, J.M. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, *2*, 57-64.

McCrae, R.R. & Costa, P.T. (1983). Social desirability scales: more substance than style. *Journal of Consulting and Clinical Psychology*, *51*, 882-888.

McDonald, R.P. (1999). Test theory: A unified approach. *Mahwah, NJ: Lawrence Earlbaum*.

McGee, R.K. (1962). The relationship between response style and personality variables: I. The measurements of response acquiescence. *The Journal of Abnormal and Social Psychology*, *64*, 229.

Messick, S., & Frederiksen, N. (1958). *Personality and ability correlates of acquiescent response set and authoritarian content*. Educational Testing Service: Princeton N.J.

Moors, G., Kieruj, N.D., & Vermunt, J.K. (2014). The effect of labeling and numbering of response scales on the likelihood of response bias. *Sociological Methodology*, *44*, 369-399.

Peabody, D. (1962). Two components in bipolar scales: Direction and extremeness. *Psychological Review*, *69*, 65-73.

Peabody, D. (1966). Authoritarianism scales and response bias. *Psychological Bulletin*, *65*, 11-23.

Rammstedt, B., Kemper, C.J., & Borg, I. (2013). Correcting Big Five Personality measurements for acquiescence: An 18-country cross-cultural study. *European Journal of Personality*, *27*, 71-81.

Ray, J.J. (1970). The development and validation of a balanced dogmatism scale. *Australian Journal of Psychology*, *22*, 253-260.

Ray, J.J. (1974). Balanced dogmatism scales. *Australian Journal of Psychology*, *26*, 9-14.

Ray, J.J. (1979). Is the dogmatism scale irreversible?. *South African Journal of Psychology*, *9*, 104-107.

Rokeach, M. (1967). Authoritarianism scales and response bias: comment on Peabody's paper. Psychological Bulletin, *67*, 349-355.

Savalei, V., & Falk, C.F. (2014). Recovering Substantive Factor Loadings in the Presence of Acquiescence Bias: A Comparison of Three Approaches. *Multivariate Behavioral Research*, *49*, 407-424.

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, *27*, 236-247.

Table 1

Abbreviated item stems for the markers (M) and content (C) items

Item	Abbreviated item content
M1	Like nothing better than having breakfast in bed
M2	Nothing is worse than an offensive odor
M3	Few things more satisfying than splurge on something
M4	Wonderful feeling to sit surrounded by your possessions
C1(F)	People cannot be blamed for some inconsistencies in what they think
C2(F)	Many problems have more than one acceptable solution
C3(D)	Careful not to compromise with those who believe differently from you in religion
C4(D)	Two kinds of people in the world: for the truth and against the truth
C5(D)	Gets enthusiastic about many causes is likely to be a “wishy-washy” person
C6(F)	Try to keep a fairly open mind in most issues
C7(F)	Unwise to indulge in generalizations. Individual circumstances often alter cases
C8(D)	Better to be a dead hero than a live coward
C9(D)	In this complicated world only way to know what’s going on is to rely on leaders/experts
C10(F)	What might be right for one person is not necessarily right for his neighbour
C11(F)	Churches don’t give enough weight to individual conscience
C12(D)	Who does not believe in some great cause has not really lived
C13(D)	Many details still remain but we have definite answers to most practical problems in life
C14(F)	There is good in everyone
C15(D)	Group which tolerates too much difference of opinion cannot exist for long
C16(F)	Possible that there are many faces to the “truth”
C17(F)	Doesn’t matter much what religion a person follows
C18(D)	Who is extremely tolerant of widely different viewpoints has few opinions on his own

Note. D: dogmatism-oriented; F: Flexibility oriented

Table 2

Calibration Results: bidimensional oblique solution in the binary and graded-response formats

Items	(a) Binary Format		(b) Graded Format	
	F1	F2	F1	F2
M1	0.00	0.40	0.00	0.33
M2	0.00	0.45	0.00	0.54
M3	0.00	0.27	0.00	0.50
M4	0.00	0.34	0.00	0.47
C1+	0.53	0.10	0.42	0.23
C2+	0.73	0.30	0.61	0.33
C3–	0.06	-0.19	0.06	-0.31
C4–	0.37	-0.12	0.32	-0.36
C5–	0.37	-0.11	0.37	-0.11
C6+	0.53	0.13	0.54	0.19
C7+	0.59	0.24	0.48	0.18
C8–	0.05	-0.18	0.12	-0.16
C9–	0.22	-0.08	0.18	-0.19
C10+	0.75	0.33	0.54	0.14
C11+	0.16	0.04	0.22	0.27
C12–	0.13	-0.19	0.16	-0.21
C13–	0.14	-0.04	0.22	-0.43
C14+	0.53	0.19	0.50	0.40
C1–	0.23	-0.36	0.42	-0.22
C16+	0.44	0.04	0.65	0.36
C17+	0.20	-0.19	0.34	0.11
C18–	0.26	-0.36	0.37	-0.27
ϕ	-0.53 (-0.63; -0.43)		-0.42 (-0.49; -0.33)	

Table 3

Reliability estimates of the content and ACQ EAP factor scores in both formats

	Content	ACQ
Binary	0.79	0.63
Graded	0.80	0.73

Table 4

Across-format correlations for the content and ACQ factor scores

	Content	ACQ
Raw	0.69 (0.66;0.72)	0.52 (0.48;0.56)
Disattenuated	0.87 (0.83;0.91)	0.73 (0.71;0.83)