J Rat-Emo Cognitive-Behav Ther (2017) 35:111–124
DOI 10.1007/s10942-016-0243-1

Assessing the Factorial Validity of the Attitudes and Belief Scale 2-Abbreviated Version: A Call for the Development a Gold Standard Method of Measuring Rational and Irrational Beliefs

Philip Hyland[1] · Robert Fox[1] · Gary Treacy[1] · Phil Maguire[2] · Daniel Boduszek[3] · Frédérique Vallières[4]

Published online: 20 May 2016

Abstract Rational emotive behaviour therapy (REBT) does not possess a measure of rational and irrational beliefs that meets internationally recognised standards for acceptable psychometric properties. Without such a measure the theory/practice of REBT cannot be rigorously evaluated, thus undermining its scientific veracity. The current study investigates the validity and reliability of a recently developed measure of rational and irrational beliefs: the Attitudes and Belief Scale 2-Abbreviated Version (ABS-2-AV). University students from three countries completed the ABS-2-AV (N = 397). An alternative models framework using confirmatory factor analysis indicated that a theoretically consistent eight-factor model of the ABS-2-AV provided the best fit of the data. A number of post hoc modifications were required in order to achieve acceptable model fit results, and these modifications revealed important methodological limitations with the ABS-2-AV. Results indicated that the validity of the ABS-2-AV was undermined due to items measuring both the psychological process of interest (rational and irrational beliefs) and the context in which these beliefs processes are presented. This is a serious methodological limitation of the ABS-2 and all questionnaires derived from it, including the ABS-2-AV. This methodological limitation resulted in the ABS-2-AV possessing poor internal reliability. These limitations are discussed in relation to the broader REBT literature and the impact such problems have on research and practice. A call is made for REBT researchers to come together to develop a "gold

✉ Philip Hyland
Philip.hyland@ncirl.ie

[1] School of Business, National College of Ireland, IFSC, Mayor Street, Dublin 1, Ireland

[2] Department of Computer Science, National University of Ireland, Maynooth, Kildare, Ireland

[3] Department of Behavioural and Social Sciences, University of Huddersfield, Huddersfield, UK

[4] Centre for Global Health, Trinity College Dublin, Dublin, Ireland

standard" method of assessing rational and irrational beliefs that meets international standard for psychometric excellence.

Keywords Attitudes and belief scale · Irrational beliefs · Rational beliefs · Rational emotive behaviour therapy (REBT) · Validity · Reliability

Introduction

Rational emotive behaviour therapy (REBT) is based on the theoretical proposition that cognitive processes (rational and irrational beliefs) mediate the impact of stressful life events on the development of cognitive, emotional, behavioural, and psychobiological responses (Hyland and Boduszek 2012). Specifically, irrational beliefs about life events are predicted to generate dysfunctional outcomes, while rational beliefs about life events are predicted to generate functional outcomes. Rational and irrational beliefs are hypothesised to reflect unique psychological processes, rather than merely representing two ends of the same psychological process (DiLorenzo et al. 2011; Dryden and David 2008). REBT theory is distinguishable from alternative cognitive–behavioural models of psychopathology on a number of grounds. One unique factor relates to the proposed nature of rational and irrational beliefs as key predictors of psychological outcomes (Hyland and Boduszek 2012).

According to REBT theory, irrational beliefs reflect four distinct cognitive processes: (1) demandingness beliefs—rigid beliefs regarding how things should be, must be, have to be etc.; (2) catastrophizing beliefs—extreme evaluations of the negative outcomes of a given event; (3) frustration intolerance (FI) beliefs—extreme negative evaluations of the capacity one has to withstand adverse life events; and (4) self-downing beliefs—extreme evaluations of one's whole self as defective in some regard. Contrastingly, rational beliefs reflect four distinctive ways of evaluating the world: (1) preference beliefs—flexible beliefs about how one would like things to be; (2) realistic evaluations of badness (REB) beliefs—non-extreme evaluations regarding the consequences of undesirable life events; (3) frustration tolerance (FT)—realistic evaluations of one's capacity to withstand adverse life events; and (4) self-acceptance beliefs—beliefs that acknowledge the fallibility of one's self and the recognition that the self is too complex to be globally rated. Numerous empirical studies have demonstrated that irrational beliefs are predictive of a range of psychopathological outcomes including depression (Szentagotai et al. 2008), posttraumatic stress symptoms (Hyland et al. 2014b), exam-related anxiety (DiLorenzo et al. 2007), and general psychological distress (Cristea et al. 2013). Additionally, studies have begun to demonstrate that rational beliefs can act as cognitive protective factors against psychopathological responses to adverse life events (Hyland et al. 2014a).

Given the central role of rational and irrational beliefs in REBT theory, the need for well-validated self-report measures of these constructs is critical. Without a valid and reliable measure of the fundamental theoretical constructs of the theory, no meaningful test of the theory can be conducted. It is disconcerting therefore that

existing self-report measures of rational and/or irrational beliefs have been heavily criticized for lacking psychometric support (Terjesen et al. 2009). Such criticism generally refers to a lack of consistency between established theory and the constructs measured (poor content validity), and use of items in existing questionnaires that measure both cognition and emotion. In their assessment of existing measures of rational and irrational beliefs, Terjesen et al. concluded that none of the available self-report measures possessed acceptable psychometric properties.

Some researchers have argued that a more recently developed, and frequently cited questionnaire, the Attitudes and Beliefs Scale-2 (ABS-2; DiGiuseppe et al. 1988), has overcome some of these psychometric limitations (e.g., Mogase et al. 2013). Unfortunately, this does not appear to be the case. In a recent study, the psychometric properties of the ABS-2 were evaluated using confirmatory factor analytic (CFA) procedures (Hyland et al. 2014c) with results indicating that the ABS-2 possessed poor factorial validity. Findings indicated that the source of the poor validity was related to the scale items being contaminated by contextual factors (items of the ABS-2 were constructed to not only measure the cognitive processes of interest but to also measure various contexts in which rational and irrational beliefs could be experienced). The authors also reported poor discriminant validity, as correlations between the cognitive processes were all very high. Hyland et al. attempted to overcome the methodological limitation of the ABS-2 by removing 'contaminated' items. Through application of confirmatory bifactor modelling procedures the authors were able to identify items from the ABS-2 that were primarily measuring the cognitive processes of interest (the rational and irrational beliefs) and were not measuring the contextual factors (e.g., comfort, affiliation, and achievement) that were the source of miss-fit. The authors thus developed a 24-item abbreviated version of the ABS-2 (the ABS-2-AV; Hyland et al. 2014c) which possessed satisfactory factorial validity and internal reliability (Cronbach's alpha). This approach was problematic however, as item selection was based on a small and unique sample of the population (trauma-exposed law-enforcement officers). Ultimately the authors merely fitted the data to the unique profile of the sample utilised in that study. There is thus little reason to think that the results of that study would generalise to alternative populations. The factorial validity and internal reliability of the ABS-2-AV is therefore still yet to be established.

Another recently developed measure of rational and irrational beliefs is the Rational and Irrational Beliefs Scale (RAIBS: Mogase et al. 2013). This scale is also undermined by the same methodological problem as the ABS-2, as items included in this scale were based on the ABS-2 (DiGiuseppe et al. 1988). The authors maintained the use of items that simultaneously measure cognitive process (rational or irrational beliefs) and the context in which these cognitive processes occur (achievement, approval, comfort, justice, and control). As Hyland et al. (2014c) empirically demonstrated, item construction in this manner means that it is difficult to determine if one is responding to these items due to the latent psychological construct of interest (the rational or irrational belief), or if one's response is being determined by the latent contextual factor in which the belief is presented. Simply put, the nature of the items included in the ABS-2 (DiGiuseppe et al. 1988) means

that it, and any measure derived from it, are very unlikely to possess satisfactory psychometric properties.

The current study is conducted in order to test the construct validity and internal reliability of the ABS-2-AV (Hyland et al. 2014c) within a distinct sample of participants from which the scale was originally developed and tested. A series of alternative, theoretically-consistent models of the factor structure of the ABS-2-AV will be compared using CFA procedures. Composite reliability analysis will be used to better establish the internal reliability of the scale as this method has been demonstrated to be superior to traditional measures of reliability such as Cronbach's alpha (Raykov 1998). Based on previously reported findings (Hyland et al. 2014c), we hypothesised that a theoretically derived eight-factor model would provide the best fit of the data but that modifications to this model would be required in order to achieve satisfactory model fit. Furthermore, given the item contamination (cognition and context) we hypothesised that the ABS-2-AV would possess unsatisfactory internal reliability.

Methods

Participants and Procedures

The sample for the current study was comprised of undergraduate psychology and computer science students recruited from four universities in the Republic of Ireland, Northern Ireland, and England ($N = 397$). The sample included a similar number of men (n = 191, 49.6 %) and women (n = 194, 50.4) with an average age of 23.33 years (SD 7.91, range 18–60). Students had spent on average 2.23 years in third-level education (SD 1.20, range 0–7) at the time of assessment. The majority of students lived in urban or suburban environments (n = 252, 65.7 %) and were single (n = 315, 82 %). All participants were selected in an opportunistic fashion and data was collected during the academic calendar from September 2013 to May 2014. Ethical permission to conduct the study was obtained from the ethical review board at the institution to which the corresponding author belongs. All students under the age 18 were excluded from the study. Participants were assured of confidentiality, instructed that they were under no obligation to participate, and could withdraw at any time. Participants completed the questionnaires using a paper-and-pencil format in their regular lecture theatre and laboratory settings. No inducements or incentives (e.g., course credit) were used to recruit volunteers.

Measures

The Attitudes and Belief Scale 2-Abbreviated Version (ABS-2-AV: Hyland et al. 2014c) is a 24-item self-report measure of rational and irrational beliefs consistent with contemporary REBT theory. The scale is intended to measure the four irrational belief processes (demandingness, catastrophizing, FI, and self-downing) and the four rational belief processes (preferences, REB, FT, and self-acceptance). Each item is scored along a five-point Likert scale from 1 ("Strongly Disagree") to

5 ("Strongly Agree"), with higher scores in each case indicating greater endorsement of a given belief process.

Analytical Plan

Five models of the latent structure of the ABS-2-AV were specified and tested using CFA procedures in Mplus version 7.11 (Muthén and Muthén 1998–2013). Robust maximum likelihood estimation was selected as this method has been demonstrated to be robust to non-normally distributed data, and to be superior to alternative methods of handling missing data such as listwise deletion (Yuan and Bentler 2000). Model 1 is a one-factor model in which responses to all 24 items are explained in terms of a single latent variable. Model 2 is a two-factor model in which the 12 irrationally-phrased items are explained by one factor ('Irrationality'), and the 12 rationally-phrased items are explained by another latent factor ('Rationality'). Model 3 is a four-factor model in which the rational and irrational belief processes are considered as two ends of a single continuum. In other words, the items purporting to measure the irrational (e.g., Demandingness) and rational (e.g., Preferences) processes load onto one factor. Model 4 is an eight-factor model in which three items load onto each of the rational and irrational belief processes. This model is consistent with the theory of REBT (David and Szentagotai 2006). Model 5 is a higher-order variant of Model 4 in which the covariation between the four irrational beliefs is explained by a higher-order 'Irrationality' latent variable, and the covariation between the four rational beliefs is explained by a higher-order 'Rationality' latent variable.

Kline's (2011) recommendations for determining model fit were followed for the CFA analyses. Good model fit is indicated by a Chi square-to-degrees of freedom (χ^2:*df*) ratio of less than 3:1. Comparative Fit Index (CFI; Bentler 1990) and Tucker Lewis Index (TLI; Tucker and Lewis 1973) values greater than .90 suggest adequate model fit, while values greater than .95 suggest excellent model fit. Values for the root-mean-square error of approximation with 90 % confidence intervals (RMSEA 90 % CI Steiger 1990) and the standardized root-mean-square residual (SRMR: Joreskog and Sorbom 1981) of .08 or less are indicative of adequate model fit, while values of .05 or less indicate excellent model fit. The Akaike Information Criterion (AIC; Akaike 1987) and Bayesian Information Criterion (BIC; Schwarz 1978) were used to evaluate alternative nested models, with the smaller value in each case indicating the best fitting model. The CFI, RMSEA, BIC, and AIC all have explicit penalties for model complexity.

Results

Model Fit Results

The model fit results are reported in full in Table 1 and, as can be seen, Models 1, 2, 3, and 5 were all found to be poor approximations of the data and were thus rejected. Consistent with our prediction, the eight-factor model of the ABS-2-AV was found

Table 1 Fit indices for the alternative factor models of the ABS-2_AV

Model	χ^2	*df*	CFI	TLI	RMSEA (90 % CI)	SRMR	AIC	BIC
Model 1	1260*	252	.538	.494	.100 (.095–.106)	.098	26,863	27,150
Model 2	1140*	251	.592	.552	.094 (.089–.100)	.091	26,709	27,000
Model 3	855*	246	.721	.687	.079 (.073–.085)	.076	26,393	26,703
Model 4	518*	224	.865	.833	.058 (.051–.064)	.056	26,039	26,437
Model 5	681*	243	.799	.772	.067 (.061–.073)	.079	26,194	26,517
Modified Model 4	397*	222	.919	.900	.045 (.038–.052)	.048	25,917	26,323

N = 397; χ^2 = Chi square goodness of fit statistic; *df* = degrees of freedom; *CFI* comparative fit index, *TLI* Tucker Lewis Index, *RMSEA* (90 % CI) root-mean-square error of approximation with 90 % CI, *SRMR* standardized square root mean residual, *AIC* Akaike Information Criterion, *BIC* Bayesian Information Criterion

* Indicates χ^2 are statistically significant ($p < .001$)

to provide the best representation of the data. Despite providing the best fit of the data, the overall model fit statistics were somewhat unsatisfactory. The χ^2:*df* ratio, RMSEA, and SRMR results suggested adequate model fit, while the CFI and TLI values were below the criteria for acceptable fit. We thus inspected the modification indices provided by Mplus to determine the source of the model misspecification. Following the methodological procedures set forth by Byrne (2011), the eight-factor model was adapted in an exploratory manner through the introduction of a model parameter that would result in the largest improvement in overall fit. This process is conducted in a sequential fashion (only one modification is made at a time) until such time as satisfactory model fit is obtained. Importantly, any modifications made to the model had to be understandable from a theoretical perspective, and were not introduced simply on statistical grounds (improvement in model fit).

Based on the modification indices, item 3, which was intended to measure FI ("It's unbearable to fail at important things, and I can't stand not succeeding at them") was shown to also be an indicator of the Demandingness factor. A cross-factor loading was introduced and this led to an improvement in fit ($\chi^2 = 417$, $df = 223$; CFI = .911; TLI = .889; RMSEA 90 % CI .047 (.040–.054); SRMR = .050). The TLI value still failed to satisfy acceptable standards and thus modification indices were again inspected. This revealed a high residual covariance between item 6 ("I do not like to be uncomfortable, tense or nervous, but I can tolerate being tense") and item 7 ("When life is hard and I feel uncomfortable, I realize it is not awful to feel uncomfortable or tense, only unfortunate and I can keep going"). While item 6 purports to measure FT beliefs and item 7 purports to measure REB beliefs, it was judged appropriate to include a residual covariance between these two items in the model given the similarity in phrasing (reference to withstanding uncomfortable situations). Following the inclusion of this residual covariance, the model fit the data satisfactorily (see Table 1) and therefore no further modifications were deemed appropriate (see Fig. 1 for final model).

The standardized factor loadings for the final, modified eight-factor model of the ABS-2-AV are presented in Table 2. All items loaded positively, and statistically significantly onto their respective latent factor. In the majority of cases these factor

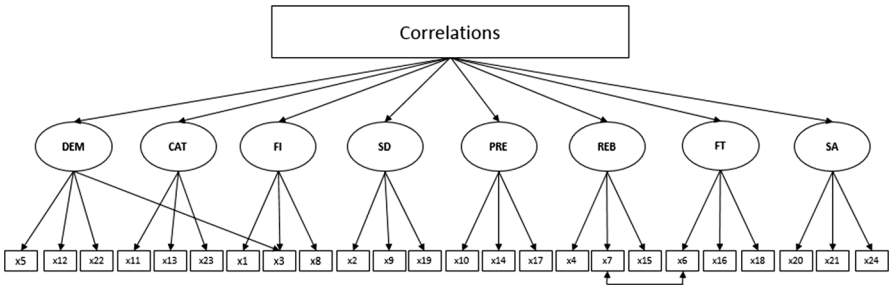

Fig. 1 Final model of the ABS-2-AV. *DEM* demandingness, *CAT* catastrophizing, *FI* frustration intolerance, *SD* self-downing, *PRE* preferences, *REB* realistic evaluations of badness, *FT* frustration intolerance, *SA* self-acceptance

loadings were of a robust magnitude, however, a small number of items possessed weak factor loadings. While item 3 loaded strongly onto the demandingness factor, it was a weak indicator of FI. Items 4 (REB) and 6 (FT) yielded factor loadings less than .30, suggesting that these items are poor indicators of their respective latent factors.

Table 3 reports the correlations between the eight factors. The four irrational belief factors were all moderately and positively associated with one another. This indicates a reasonable degree of discrimination between the factors. The four rational belief factors were also positively associated with one another. However, in this case the associations were stronger and thus indicative of poor discrimination between the rational belief factors. In particular, the FT and REB factors were very strongly associated ($r = .86$). The correlations between the rational and irrational counterparts of Demandingness and Preferences ($r = -.46$), Catastrophizing and REB ($r = -.54$), and FI and FT ($r = -.49$) were all negative and of a moderate magnitude. This suggests that these rational and irrational beliefs are predominately distinct psychological processes. Contrastingly, self-acceptance and self-downing beliefs ($r = -.76$) were strongly (negatively) associated with one another, suggesting that these rational-irrational beliefs may be more reflective of the same underlying psychological process that then other belief processes.

Composite Reliability Analysis

The use of traditional measures of internal reliability such as Cronbach's alpha have been criticised within a latent variable modelling context given the propensity to over- or under-estimate scale reliability (Raykov 1998). In order to provide a more rigorous assessment of the internal reliability of the ABS-2-AV, the current study estimated the composite reliability of the measurement properties of the scale. Values greater than .60 are considered acceptable (Bagozzi and Yi 1988; Diamantopoulos and Winklhofer 2001). The reliability of demandingness (including item 3; .75), catastrophizing (.66); self-downing (.76), preferences (.76), and self-acceptance (.78) were all acceptable. The composite reliability for FI (.58) and FT

Table 2 Standardized factor loadings (and standard errors) for the modified eight-factor model of the ABS2_AV

Item	β	SE
Demandingness		
5. I must do well at important things, and I will not accept it if I do not do well	.71	.05
12. It's essential to do well at important jobs; so I must do well at these things	.61	.05
22. I must be successful at things that I believe are important, and I will not accept anything less than success	.64	.05
3. It's unbearable to fail at important things, and I can't stand not succeeding at them	.65	.06
Catastrophizing		
11. It's awful to be disliked by people who are important to me, and it is a catastrophe if they don't like me	.58	.06
13. Sometimes I think the hassles and frustrations of everyday life are awful and the worst part of my life	.69	.04
23. If loved ones or friends reject me, it is not only bad, but the worst possible thing that could happen to me	.61	.05
Frustration intolerance		
1. Its unbearable being uncomfortable, tense or nervous and I can't stand it when I am	.59	.06
3. It's unbearable to fail at important things, and I can't stand not succeeding at them	.16	.06
8. I can't stand being tense or nervous and I think tension is unbearable	.82	.05
Self-downing		
2. If important people dislike me, it is because I am an unlikable bad person	.61	.05
9. If I do not perform well at tasks that are very important to me, it is because I am a worthless bad person	.74	.05
19. When people I like reject me or dislike me, it is because I am a bad or worthless person	.79	.04
Preferences		
10. I do not want to fail at important tasks but I realize that I do not have to perform well just because I want to	.62	.05
14. I want to perform well at some things, but I do not have to do well just because I want to	.68	.05
17. I want to do well at important tasks, but I realize that I don't have to do well at these important tasks just because I want to	.77	.04
Realistic evaluation of badness		
4. It is unfortunate when I am frustrated by hassles in my life, but I realize it's only disappointing and not awful to experience hassles	.22*	.10
7. When life is hard and I feel uncomfortable, I realize it is not awful to feel uncomfortable or tense, only unfortunate and I can keep going	.46	.08
15. It's bad to be disliked by certain people, but I realize it is only unfortunate to be disliked by them	.42	.07
Frustration tolerance		
6. I do not like to be uncomfortable, tense or nervous, but I can tolerate being tense	.29	.07
16. I get distressed if I'm not doing well at important tasks, but I can stand the distress of failing at important tasks	.54	.06
18. It's only frustrating not doing well at some tasks, but I know I can stand the frustration of performing less than well	.42	.07
Self-acceptance		
20. When people whom I want to like me disapprove of me, I know I am still a worthwhile person	.61	.05

Table 2 continued

Item	β	SE
21. Even when my life is tough and difficult, I realize that I am a person who is just as good as anyone else even though I have hassles	.85	.03
24. When my life becomes uncomfortable, I realize that I am still a good person even though I am uncomfortable	.74	.04

All Factor loadings are statistically significant ($p < .001$), except * ($p = .02$)

Table 3 Correlations (standard errors) between the eight factors of the ABS-2-AV

	1	2	3	4	5	6	7	8
1. Demandingness	–							
2. Catastrophizing	.48 (.07)	–						
3. Frustration intolerance	.40 (.07)	.66 (.07)	–					
4. Self-downing	.29 (.06)	.61 (.06)	.33 (.07)	–				
5. Preferences	−.46 (.07)	−.30 (.08)	−.22 (.07)*	−.30 (.07)	–			
6. Realistic evaluation of badness	−.27 (.11)**	−.54 (.11)	−.55 (.11)	−.66 (.12)	.69 (.14)	–		
7. Frustration tolerance	−.52 (.08)	−.50 (.08)	−.49 (.09)	−.41 (.08)	.77 (.06)	.86 (.12)	–	
8. Self-acceptance	−.15 (.07)**	−.50 (.07)	−.24 (.07)*	−.76 (.04)	.29 (.07)	.72 (.11)	.56 (.07)	–

All factor correlations are statistically significant ($p < .001$), except * ($p < .01$), and ** ($p < .05$)

(.53) were slightly below satisfactory levels, while the REB scale (.32) possessed poor reliability.

Discussion

The field of REBT has been criticised for a lack of scientific rigor in the testing of its theoretical hypotheses and treatment efficacy (Padesky and Beck 2003). A fundamental perquisite for such research to be undertaken is the availability of a reliable and valid method of measuring the central theoretical constructs (rational and irrational beliefs). The field of REBT has suffered greatly from the lack of a "gold standard" measure of rational and irrational beliefs that possesses excellent psychometric properties, is congruent with contemporary theory, and is consistently used among various researchers in the field so that results can be easily compared and collated (Terjesen et al. 2009).

Results of the current study were both encouraging and discouraging for the developing science of REBT. On a positive note, the theoretically consistent eight-factor model of the ABS-2-AV was found to provide superior fit to the alternative models tested. Together with previous factor analytic data (Hyland et al. 2014c)

there appears to be consistent support for the basic structure of rational and irrational beliefs. However, this positive finding must be viewed in light of the fact that the actual fit of the model to the obtained data unsatisfactory. In an absolute sense, the REBT model of rational and irrational beliefs failed to adequately explain the nature of responses to the 24 items of the ABS-2-AV. This finding suggests that while the underlying theory may be accurate, the method of assessing this theory is problematic.

The source of the measurement problem is related to the inability of certain items to accurately measure the psychological construct that they are intended to measure. One item specified to measure FI beliefs was found to be a measure of Demandingness beliefs, while two items, one intended to measure REB beliefs and the other intended to measure FT beliefs, were found to possess a high degree of similarity. The similarity in responses to these two items was due to some unmeasured latent variable. Inspection of the content of these two items revealed that both items related to the belief in one's ability to manage uncomfortable situations. This result is indicative of the major methodological limitation associated with the general ABS-2 questionnaire, namely that the indicators included in these scales contain measures of both cognitive processes (rational and irrational beliefs) and context (the context in which the belief may manifest); and this finding is consistent with the results of Hyland et al. 2014c. The simultaneous measurement of the belief processes and the context in which they may occur will negatively affect the measurement properties of any scales that includes such items.

Further problems with the ABS-2-AV were evident through the inspection of the factor loadings, factor correlations, and composite reliability results. Overall, the findings revealed that the ABS-2-AV provided a reasonably acceptable (although far from ideal) measure of the irrational belief processes, and a poor measure of the rational belief processes. With the exception of the cross-factor loading problem with item 3 (measuring Demandingness rather than Catastrophizing), all items were found to be reasonably good measures of their intended irrational belief process. Internal reliability, determined via the use of composite reliability analysis, also signified a good degree of consistency among the items of each factor. The correlations between the four irrational belief processes ranged from weak ($r = .29$ between demandingness and self-downing) to strong ($r = .66$ between LFT and catastrophizing), suggesting reasonable levels of discrimination between the factors. Issues of multicollinearity in predictive modelling studies should therefore not be of major concern.

A different picture was evident in relation to the measurement of the rational belief processes. Many of the items used to measure REB and FT beliefs possessed very low factor loadings, indicating that these items are poor measures of their intended rational belief processes. Consequently the items used to measures these factors possessed poor internal reliability, particularly in the case of the REB factor. In addition, the correlations between the four rational belief processes were generally strong (with the sole exception of the relationship between preferences and self-acceptance), indicating poor discrimination between the factors. The lack of satisfactory internal reliability and discrimination between the factors means that predictive modelling using the rational belief scores becomes highly problematic.

This is particularly undesirable given the extremely limited data that currently exists regarding the role of rational beliefs in mitigating psychopathological responses and in promoting psychologically healthy outcomes (DiLorenzo et al. 2011; Hyland et al. 2014a).

Current results were interesting to consider in light of the REBT theoretical proposition that rational and irrational beliefs are not bipolar psychological constructs, but rather reflect two separate continua (DiLorenzo et al. 2011; Dryden and David 2008). The results were generally supportive of this proposition, as the correlations between demandingness and preferences, catastrophizing and REB, and FI and FT were all of a moderate nature, suggesting that high scores on one belief process are not necessarily reflective of low scores on the other. Current results are therefore consistent with previous findings regarding the relationship between these rational and irrational beliefs (e.g., DiLorenzo et al. 2011; Fulop 2007; Macavei 2005). The exception to this pattern was the observed relationship between self-downing and self-acceptance beliefs. The strength of the association between these beliefs indicates that those with high levels of self-downing beliefs are likely to possess low levels of self-acceptance beliefs, and vice versa. This finding is consistent with those reported by Hyland et al. (2014c) and suggest that self-downing and self-acceptance beliefs likely reflects separate ends of the same underlying psychological process.

These findings have significant implications for both researchers and clinicians. The field of REBT research and practice faces a serious problem given the apparent absence of a method of assessing the central theoretical constructs in a psychometrically sound manner. Without such a measure researchers are limited in their ability to rigorously test theoretical predictions, develop and advance current theory, compare the predictions of REBT theory to alternative and competing models of psychopathology, and to assess the proposed mechanism of recovery in therapeutic interventions. It is essential therefore that researchers in the field of REBT come together to develop a measure of rational and irrational beliefs that satisfies internationally recognised standards for validity and reliability (see Terwee et al. 2007). The most widely recognised criteria for establishing the psychometric properties of a given scale are provided by the Scientific Advisory Committee (SAC) of the Medical Outcomes Trust (Aaronson et al. 2002). Eight criteria are presented which include: (1) conceptual and measurement models, (2) validity, (3) reliability, (4) responsiveness, (5) interpretability, (6) respondent and administrative burden, (7) alternative forms, and (8) cultural and language adaptations (translations). The development of new measures of rational and irrational beliefs should be sensitive to each of these criteria.

It is recommended that the development of new, psychometrically sound measures of rational and irrational beliefs follow many of the same recommendations previously provided by Terjesen et al. (2009). We additionally propose that the foundation of a new "gold standard" method of measuring rational and irrational beliefs should start from the recognition of an eight factor structure (measurement of the four irrational beliefs and the four rational beliefs). Item construction must be conducted in a manner that ensures that only the cognitive process (i.e., each belief type) is measured. Every effort should be made to avoid contaminating the items of

the scale with statements that (implicitly or explicitly) measure affect, behaviour, or context. One exception to this recommendation could relate to the development of situationally-specific measures of rational and irrational beliefs. From a standard generalized measure of rational and irrational beliefs, situationally-specific measures could be derived as per the nature of the study being undertaken (e.g., a study of exam-related anxiety). In such cases, the context is explicitly identifiable within each item and relates directly to the nature of the psychological process under investigation in a given study. Furthermore, these items should be short, easy-to-understand statements that avoid confusion and response acquiescence. This will facilitate the development of translated versions of the scale that are valid across multiple cultures and contexts. Future scale developers should also be cognizant of the number of items used to measure each belief process. Single indicator items should be avoided, as they preclude assessments of reliability and factor discrimination. Nevertheless, scale developers should also avoid numerous indicators as such measures can become overly lengthy and impractical for research and clinical work. We recommend 3–5 items to measure each of the eight factors. Additionally, each item should ideally be measured along a continuous scale of agreement/disagreement (e.g., a 1–10 scale) rather than measured with a traditional ordered-categorical Likert scales.

The current study is limited due to the use of a relatively small and distinctive sample of the population. Important claims regarding the psychometric properties of a given scale would preferably be made based on a more representative sample of the population. It is possible that the current results are influenced by cultural or demographic factors related to the current cohort. However, given that current findings are consistent with previous discoveries (Hyland et al. 2014c; Terjesen et al. 2009) it is unlikely that these results are due to sampling biases. A further notable limitation with the ABS-2-AV, and all other scales derived from the ABS-2, is the sole focus on self-downing and self-acceptance beliefs, while REBT theory is explicit in stating that these global evaluative beliefs (rational and irrational) can be additionally directed towards others and the world/environment. Methods of measuring these psychological processes must also be considered in future scale developments.

In conclusion, the current study provides additional evidence to the existing REBT literature that current measures of rational and irrational beliefs are unsatisfactory. Moreover, the current study indicates that attempts to develop psychometrically sound scales from existing and widely used measures (e.g., ABS-2) are flawed given the inherent methodological limitations associated with these scales. In order to substantially advance the science and practice of REBT we call on researchers throughout the field to come together to develop a new measure of rational and irrational beliefs. Such a measure should be constructed and assessed with recognition of the necessary criteria for the establishment of psychometric excellence (Aaronson et al. 2002; Terwee et al. 2007). The joint efforts of researchers from multiple nations working in unison to develop a "gold standard" measure of rational and irrational beliefs would signify a major advancement in the field of REBT research and practice.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Aaronson, N., Alonso, J., Burman, A., Lohr, K. N., Patrick, D. L., Perrin, E., & Stein, R. E. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research, 11*, 193–205.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*, 317–332.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science, 16*, 74–94.

Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin, 107*, 238–246.

Byrne, B. M. (2011). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. London: Routledge.

Cristea, I. A., Montgomery, G. H., Szamoskozi, S., & David, D. (2013). Key constructs in "classical" and "new wave" cognitive behavioral psychotherapies: Relationships among each other and with emotional distress. *Journal of Clinical Psychology, 69*, 584–599.

David, D., & Szentagotai, A. (2006). Cognition in cognitive-behavioral psychotherapies (CBT): Toward an integrative model. *Clinical Psychology Review, 3*, 284–298.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*, 269–277.

DiGiuseppe, R., Leaf, R., Exner, T., & Robin, M. V. (1988). *The development of a measure of rational/irrational thinking.* Paper presented at the World Congress of Behavior Therapy, Edinburgh, Scotland.

DiLorenzo, T. A., David, D., & Montgomery, G. H. (2007). The interrelations between irrational cognitive processes and distress in stressful academic settings. *Personality and Individual Differences, 42*, 765–777.

DiLorenzo, T. A., David, D., & Montgomery, G. H. (2011). The impact of general and specific rational and irrational beliefs on exam distress; A further investigation of the binary model of emotional distress as an emotional regulation model. *Journal of Cognitive and Behavioural Psychotherapies, 11*, 121–142.

Dryden, W., & David, D. (2008). Rational emotive behavior therapy: Current status. *Journal of Cognitive Psychotherapy: An International Quarterly, 22*, 195–209.

Fulop, I. E. (2007). A confirmatory factor analysis of the attitude and belief scale 2. *Journal of Cognitive and Behavioral Psychotherapies, 7*, 159–170.

Hyland, P., & Boduszek, D. (2012). Resolving a difference between cognitive therapy and rational emotive behaviour therapy: Towards the development of an integrated CBT model of psychopathology. *Mental Health Review Journal, 17*, 104–117.

Hyland, P., Maguire, R., Shevlin, M., & Boduszek, D. (2014a). Rational beliefs as cognitive protective factors in posttraumatic stress responses. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 32*, 297–312.

Hyland, P., Shevlin, M., Adamson, G., & Boduszek, D. (2014b). The organisation of irrational beliefs in posttraumatic stress symptomology: Testing the predictions of REBT theory using structural equation modelling. *Journal of Clinical Psychology, 70*, 48–59.

Hyland, P., Shevlin, M., Adamson, G., & Boduszek, D. (2014c). Modelling the structure of the Attitudes and Belief Scale 2: Toward the development of an abbreviated version. *Cognitive Behaviour Therapy, 43*, 60–71.

Joreskog, K., & Sorbom, D. (1981). *LISREL V: Analysis of linear structural relationships by the method of maximum likelihood.* Chicago: National Educational Resources.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). London: The Guilford Press.

Macavei, B. (2005). The role of irrational beliefs in the rational emotive behavior theory of depression. *Journal of Cognitive and Behavioral Psychotherapies, 5*, 73–83.

Mogase, C., Stefan, S., & David, D. (2013). How do we measure rational and irrational beliefs? The development of Rational and Irrational Beliefs Scale (RAIBS)—a new theory driven measure. *Journal of Cognitive and Behavioral Psychotherapies, 13*, 529–546.

Muthén, L. K., & Muthén, B. O. (1998–2013). *MPlus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Padesky, C. A., & Beck, A. T. (2003). Science and philosophy: Comparison of cognitive therapy and rational emotive behaviour therapy. *Journal of Cognitive Psychotherapy, 17*, 211–224.

Raykov, T. (1998). Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement, 22*, 375–385.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioural Research, 25*, 173–180.

Szentagotai, A., David, D., Lupu, V., & Cosman, D. (2008). Rational-emotive behaviour therapy versus cognitive therapy versus medication in the treatment of major depressive disorder: Mechanisms of change analysis. *Psychotherapy Theory, Research, Practice, 4*, 523–538.

Terjesen, M. D., Salhany, J., & Sciutto, M. J. (2009). A psychometric review of measures of irrational beliefs: Implications for psychotherapy. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 27*, 83–96.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*, 34–42.

Tucker, L. R., & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1–10.

Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology, 30*, 165–200.