Journal of Personality and Social Psychology

Reconceptualizing Anxiety as a Continuum That Ranges From High Calmness to High Anxiety: The Joint Importance of Reducing Distress and Increasing Well-Being

Andy P. Siddaway, Peter J. Taylor, and Alex M. Wood

Online First Publication, March 2, 2017. http://dx.doi.org/10.1037/pspp0000128

Journal of Personality and Social Psychology

0022-3514/17/$12.00 http://dx.doi.org/10.1037/pspp0000128

Reconceptualizing Anxiety as a Continuum That Ranges From High Calmness to High Anxiety: The Joint Importance of Reducing Distress and Increasing Well-Being

Andy P. Siddaway
University of Stirling

Peter J. Taylor
University of Manchester

Alex M. Wood
University of Stirling

We first replicate a study by Vautier and Pohl (2009), who used the State–Trait Anxiety Inventory (STAI) to reexamine the structure of anxiety. Using two large samples ($N = 4{,}138$ and 1,824), we also find that state and trait anxiety measure continua that range from high calmness to high anxiety. We then significantly extend previous findings and make the clinical importance of this topic more explicit by characterizing the (linear or nonlinear) form of the relationship between the calmness–anxiety continuum and other psychiatric variables for the first time. This form is critical to understanding anxiety problems, as discontinuities in relationships with other psychological conditions could be used to define a natural boundary of problematic anxiety. Baseline levels on the calmness–anxiety continuum are found to have a near linear relationship with changes in depression, aggression, and substance misuse over time. Taken together, these results indicate the joint importance and usefulness of treating anxiety problems and promoting calmness, as doing so may promote resilience from developing other psychiatric conditions. Psychiatric and psychological interventions that are grounded in this continuum conceptualization would logically be stopped when an individual reports experiencing high levels of calmness. Our results point to the usefulness of early intervention and prevention (when people begin to move away from high calmness) and instilling resilience (by providing interventions to move people toward high calmness).

Keywords: anxiety, psychiatric, relax, structure, well-being

Supplemental materials: http://dx.doi.org/10.1037/pspp0000128.supp

The State-Trait Anxiety Inventory (STAI) is "the most widely used device to measure anxiety across cultures" (Lonner & Ibrahim, 1989: p. 317), and there is extensive support for its psychometric properties (see Spielberger, 1989; Spielberger & Diaz-Guerrero, 1986; Spielberger, Gorsuch, & Lushene, 1970). Strong consensus over decades of research has supported a four factor structure that consists of state and trait "anxiety present" and "anxiety absent" factors (Bernstein & Eveland, 1982; Hishinuma, Miyamoto, Nishimura, & Nahulu, 2000; Mook, Van der Ploeg, & Kleijn, 1992; Spielberger, Vagg, Barker, Donham, & Westberry, 1980; Suzuki, Tsukamoto, & Abe, 2000; Vagg, Spielberger, & O'Hearn, 1980; Vigneau & Cormier, 2008). "Anxiety absent" items (e.g., "I feel calm") are reverse scored. As currently defined, a total score on the state and trait scales therefore involves a combination of the "presence" and "absence" of anxiety symptoms.

Although there has been extensive support for a four factor structure to the STAI, there are both conceptual and statistical

Andy P. Siddaway, Behavioural Science Centre, Stirling Management School, University of Stirling; Peter J. Taylor, Division of Psychology & Mental Health, University of Manchester; Alex M. Wood, Behavioural Science Centre, Stirling Management School, University of Stirling.

Ethical approval to use two existing data sets was granted to Andy P. Siddaway by the University of Hertfordshire, United Kingdom. The data sets used also received ethical approval when they were conducted. This report is independent research arising from a Medical Research Council-funded Clinical Research Training Fellowship (Grant reference: MR/L017938/1) awarded to Andy P. Siddaway. The views expressed in this publication are those of the Fellow and not necessarily those of the MRC. The MRC had no role in the writing or submission of this article.

We are grateful to Professor Green for granting permission to use the Cambridge Prenatal Screening Study data and to the United Kingdom Data Archive. We are also grateful to the National Center on Indigenous Hawaiian Behavioral Health (NCIHBH) for the Native Hawaiian data set. The NCIHBH study was supported by the Research Centers in Minority Institutions (RCMI) Supplement, National Institutes of Health (NIH; RR0361-06S1), and the National Institute of Mental Health (NIMH; R24 MH5015-01, R24 MH57079-A1). The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the funding agencies.

Correspondence concerning this article should be addressed to Andy P. Siddaway, Behavioural Science Centre, Stirling Management School, 3B55 Cottrell Building, University of Stirling, Stirling, FK9 4LA, Scotland. E-mail: andy.siddaway@stir.ac.uk

reasons to suggest that this issue needs to be revisited, with implications for how we understand the structure of anxiety. The conceptual argument was provided by Joseph and Wood (2010), who observed that reverse scored STAI items appear to assess calmness and relaxation (e.g., "I am cool, calm and collected" [trait anxiety]; "I feel calm" [state anxiety]), rather than merely the absence of anxiety problems. They pointed out that for the lowest possible score on the STAI to occur, a person would have to give all of the positively worded items (e.g., "I feel calm") the highest possible score ("Almost always") *and* all of the negatively worded items (e.g., "I feel anxious") the lowest possible score ("Almost never"). On this basis, they suggested that the lowest score on the STAI does not just indicate the *absence* of anxiety problems, as has been convention in research and practice for decades; it actually indicates the *presence* of calmness and relaxation. The STAI, as conventionally coded, could therefore be (re)conceptualized as ranging from high calmness to high anxiety. Although this interpretation of the STAI is new, the idea that anxiety might form a continuum with calmness has previously been recognized by the circumplex model of affect (e.g., Kuppens, Tuerlinckx, Russell, & Barrett, 2013; Russell, 1980; Russell & Carroll, 1999; Yik, Russell, & Steiger, 2011) and by Carver and Scheier's (1998) model of affect.[1]

The statistical argument for reexamining the factor structure of the STAI is as follows. It is common practice for self-report psychiatric scales to include some positively worded items that are reverse scored to compute a total score (Woods, 2006). The rationale is that these items measure the same construct as the negatively worded items (psychiatric symptoms), while reducing the tendency for respondents to agree more than disagree (acquiescence bias), or respond according to their general feeling about the topic rather than the specific content of the items (a response set; Green, Goldman, & Salovey, 1993; Woods, 2006). However, there is evidence that including positively worded items in psychiatric scales can inadvertently lead to the existence of a separate "method factor" that is not substantively meaningful (Green et al., 1993; Vautier & Pohl, 2009; Woods, 2006). Two Monte Carlo studies demonstrate that the existence of two factors would be inferred from the normal methods of exploratory and confirmatory factor analysis if only 10% of respondents respond carelessly to positively worded items (Schmitt & Stults, 1985; Woods, 2006). When separate factors within a scale respectively contain only positively worded and negatively worded items, the potential substantial importance of each factor is confounded with potential artifactual effects (Woods, 2006). These observations have led to a growing consensus among methodologists that factor analytic models need to account for item wording when demonstrating the existence of separate substantive factors in scales that include reverse scored items (Woods, 2006).

One study to date has examined the factor structure of the STAI while accounting for the potential influence of positively worded items (Vautier & Pohl, 2009). The authors used structural equation modeling (SEM) to control for item wording (method effects) in the French adaptation of the STAI. Trait and state STAI scales were each found to measure one construct rather than separate "anxiety present" and "anxiety absent" factors (Vautier & Pohl, 2009).

Although these structural findings have important implications for the scoring and interpretation of the STAI, the broad applicability and clinical significance of these findings was limited in at least two important respects. First, this study utilized an exclusively adult sample who completed the French adaptation of the STAI. It is therefore unclear at present whether these results would generalize beyond adults and to the original and more commonly used English language version of the scale.

Second, the form of the relationship between the calmness–anxiety continuum and other psychiatric variables remains to be established. This form is critical to understanding the boundaries of anxiety problems, as it may be that anxiety has a stronger relationship with other difficulties when it reaches a particular level, thereby demarcating what is psychopathological versus what is not (Markon, 2010). Flett, Vredenburg, and Krames (1997) referred to this as "phenomenological continuity," that is, continuity in the relationship between psychopathology and its antecedents, concomitants, or sequalae. Thus, even if anxiety is relatively continuous in a psychometric sense, its relationship with associated variables might be relatively discontinuous or nonlinear in form, thereby defining a natural boundary of problematic anxiety (Markon, 2010).

The Current Study

Given the potential practical importance of understanding the structure of anxiety, the present study sought to replicate and extend the findings of Vautier and Pohl (2009). We begin by using two large, diverse samples to test whether a hypothesized continuum is apparent in the most commonly used, English language version of the STAI when the influence of positively worded items is accounted for using SEM. We then extend the findings of Vautier and Pohl (2009) by characterizing the (linear or nonlinear) form of the relationship between the calmness–anxiety continuum and other psychiatric variables for the first time. This approach parallels work on stress, for example, which has established an inverted-U-shaped relationship between stress and memory function in that memory performance is impaired under conditions above or below optimal stress levels (Broadbent, 1965; Salehi, Cordero, & Sandi, 2010; Yerkes & Dodson, 1908). These analyses make the clinical importance of this topic more apparent than merely examining whether a continuum exists. An analysis of related variables in several domains is required to establish phenomenological continuity (Flett et al., 1997); we use measures of depression, substance abuse, and aggression, as these are often comorbid with anxiety (Kendler, Prescott, Myers, & Neale, 2003; Mineka, Watson, & Clark, 1998). We predict that moving from anxiety to calmness will provide an equal (near linear) decrease in risk for several other psychiatric variables over time, irrespective of position on the calmness–anxiety continuum (total score on the STAI).

Clarifying the form of the relationship between the calmness–anxiety continuum and other psychiatric variables has potentially significant implications for the conceptualization of psychological problems and clinical practice. For example, it is possible that

[1] Carver and Scheier (1998) suggest that the avoidance system underlies this continuum and that calmness or anxiety are experienced depending on an individual's perceived effectiveness (or ineffectiveness) in moving away from threat.

there is no relationship between the calmness–anxiety continuum and other psychiatric variables up to a particular point (e.g., throughout the range of the calmness pole), after which anxiety symptoms come to have an increasingly detrimental effect. Evidence in support of this conceptualization would point to the well-known deleterious effects of anxiety symptoms and therefore support the current emphasis in mental health services on alleviating and treating anxiety problems. This conceptualization of anxiety (and other mental health) problems underpins psychiatric nomenclature and, as a result, psychiatric and psychological interventions tend to be stopped at the point of anxiety problem absence.

An important alternative conceptualization is one in which the relationship between the calmness–anxiety continuum and other psychiatric variables is linear throughout the range of the continuum. Such a relationship would be consistent with increasing discussions regarding whether or not it could be advantageous (or cost-effective) for clinical services to promote well-being. For example, some professional bodies have called for mental health professionals to jointly focus on increasing well-being and reducing distress (e.g., The British Psychological Society, 2010). Positive Clinical Psychology has endorsed a balanced and equally weighted focus on the positive and negative aspect of life (Wood & Tarrier, 2010, as clarified in Johnson & Wood, in press; Wood & Johnson, 2016). Evidence of a constant linear relationship between the calmness–anxiety continuum and other psychiatric variables would indicate both psychometric and phenomenological continuity and would therefore provide much-needed evidence to substantiate these theoretical arguments. Such evidence would point to the joint importance and usefulness of treating anxiety problems and promoting calmness since calmness and anxiety reside on the same continuum and changing one therefore changes the other. Psychiatric and psychological interventions that are grounded in this conceptualization would logically be stopped when an individual reports high calmness and could be initiated when an individual begins moving away from high calmness.

Method

Participants

Two samples were used. Sample 1 comprised 4,138 adolescents aged 13 to 21 years from Hawai'i. These individuals took part in the 5-year longitudinal Hawaiian High Schools Health Survey (HHSHS) study conducted by the National Center on Indigenous Hawaiian Behavioral Health (NCIHBH). This sample provides a broad spread of ages, ethnicities, socioeconomic status, and gender (Andrade et al., 2006; Hishinuma et al., 2001). Participants for the HHSHS study were sampled from five high schools during five consecutive school years (1991/1992 to 1995/1996). The schools were selected from both urban and rural areas to obtain a representative sample of adolescents residing in Hawaii. Students who provided assent completed the survey in their classrooms under the supervision of their teachers. Parents of students younger than 18 years old were notified of the study by mail and given an opportunity to refuse participation. Data collected during the 1992/1993 ($N = 4{,}164$), 1993/1994 ($N = 4{,}182$), and 1994/1995 ($N = 1{,}433$) school years were used in this study. These three school years were used because of the large sample size and inclusion of the variables of interest. This sample completed the STAI and measures of depression, aggression, and substance misuse. There was some missing demographic information and incomplete questionnaire responses (see Andrade et al., 2006; Hishinuma et al., 2001). We multiply imputed missing data as best practice (discussed below). However, the dataset authors report that there were no significant differences on the STAI between participants who had intact versus missing data (Hishinuma et al., 2001).

Sample 2 comprised 1,824 British pregnant women aged 16–43 years ($M = 26.8$ years) from the Cambridge Prenatal Screening Study (CPSS). The purpose of the study was to examine the knowledge, attitudes, anxieties, and experiences of pregnant women booked for antenatal care at hospitals with differing screening policies. Participants were recruited from nine District hospitals within 60 miles of Cambridge (United Kingdom; U.K.) between 1990 and 1991. They provided information by telephone and mail interviews. The response rate was 53%. This sample only completed the STAI and was included to replicate our structural analyses. Trait STAI items were completed at 12 weeks pregnant; State STAI items were completed at 35 weeks pregnant and 6 weeks postnatal. Full demographic details and procedures are reported elsewhere (Green, Statham, & Snowdon, 1996).

Measures

State–Trait Anxiety Inventory (STAI; Spielberger et al., 1970). The STAI is described as a measure of anxiety problems. It consists of 40 items. Trait anxiety is seen as a relatively stable individual difference in the tendency to respond to situations perceived as threatening with elevation in state anxiety (Spielberger et al., 1970). State anxiety is conceptualized within the STAI as a transitory emotional state characterized by subjective, consciously perceived feelings of tension and apprehension and heightened autonomic arousal. The trait scale comprises thirteen negatively worded items (e.g., "I worry too much over something that really doesn't matter") and seven positively worded items (e.g., "I am cool, calm and collected"). Trait items are rated on a 4-point frequency scale based on "how you generally feel." The state scale comprises ten negatively worded items (e.g., "I feel anxious") and ten positively worded items (e.g., "I feel calm"). State items are rated on a 4-point frequency scale, with instructions asking readers to rate based on "how you feel right now, that is, at this moment." State and trait scales demonstrate excellent internal consistency (average Cronbach's $\alpha\text{s} > .89$) and the trait scale has evidenced excellent test—retest reliability at multiple time intervals (average $r = .88$; Barnes, Harp, & Jung, 2002). The current samples demonstrated similar internal consistency values (Cronbach's alpha = .89-.90). The state scale demonstrates lower temporal stability (average $r = .70$), as would be expected given the nature of the construct (Barnes et al., 2002). Trait and state scales have evidenced adequate convergent validity with other measures of state and trait anxiety (Spielberger, 1989) and discriminant validity from, for example, aggression ($r = .38$) and substance use ($r = .19$; Knight, Waal-Manning, & Spears, 1983). The STAI has been validated with a range of ethnic groups (e.g., Boeke, Duivenvoorden, & Bonke, 1984; Canals, Marti-Henneberg, Fernandez-Ballart, Clivillle, & Domenech, 1992; Vautier & Pohl, 2009).

Centre for Epidemiological Studies-Depression (CES-D; Radloff, 1977). The CES-D is one of the most frequently used self-report measures of depressive experiences (Santor, Gregus, & Welch, 2006). Responses capture the frequency of feelings and behaviors over the past seven days and are rated on a 4-point scale ranging from 0 (*rarely or none of the time*) to 3 (*most or all of the time*). The CES-D contains twenty items that are summed so that scores have a potential range from 0 to 60, with higher scores indicating greater frequency of depressive experiences (Radloff, 1977). The CES-D has been shown to have good psychometric properties, including high internal consistency in community and psychiatric populations (Cronbach's αs = .85 - .90; Ensel, 1986; Radloff, 1977; Roberts, 1980); convergent validity with other popular measures of depressive experiences such as the Patient Health Questionnaire-9 ($r = .85$; Amtmann et al., 2014) and Beck Depression Inventory-II ($r = .86$; Shean & Baldwin, 2008); and discriminant validity from, for example, aggression ($r = .44$) and substance use ($r = .24$; Makini et al., 1996). The CES-D demonstrated excellent internal consistency in the current sample (Cronbach's alpha = .88). A cutoff score of 16 has been found to have sensitivity and specificity rates of 86.7 and 76.6 for identifying depressed individuals, whereas a cutoff score of 21 has a sensitivity and specificity rate of 73.0 and 96.1 (Shean & Baldwin, 2008). The CES-D has been validated with a range of ethnic groups (e.g., Andrade et al., 2006; Garrison, Addy, Jackson, McKeown, & Waller, 1991).

Substance Abuse Subtle Screening Inventory—Adolescent version (SASSI-A; Miller, 1990). Six items were administered from the Substance Abuse Subtle Screening Inventory-Adolescent (SASSI-A; Miller, 1990) as a brief screen for substance use and impairment and dependency arising from substance use. The SASSI-A has been shown to have good psychometric properties, including acceptable internal consistency in the current sample (Cronbach's alpha = .74); and discriminant validity from anxiety ($r = .19$), depression ($r = .24$), and aggression ($r = .33$; Makini et al., 1996). The SASSI-A also been shown to concord with a diagnosis of substance abuse and dependency on the Diagnostic Interview Schedule for Children (Nishimura et al., 2001), predict counselor *DSM–III* diagnoses for dually diagnosed adolescent inpatients (Piazza, 1996), and predict adolescent chemical dependency (Risberg, Stevens, & Graybill, 1995). The SASSI-A has been validated with a range of ethnic groups (Nishimura et al., 2001).

Braver Aggressiveness Dimension Scale (BADS; Braver, Fogas, Sandler, & Volchik, 1986). The BADS is a 14-item abbreviated self-report measure of child and adolescent aggression. It was derived from the longer Youth Self-Report scales (YSR: Achenbach, 1991), the self-report version of the Child Behavior Checklist. Items selected for the BADS were those items from the YSR which were significantly more likely to be endorsed by clinically diagnosed, conduct-disordered children and adolescents. The BADS has good psychometric properties, including good internal consistency in the current sample (Cronbach's alpha = .85); one year test-retest stability ($r = .61$); and discriminant validity from anxiety ($r = .38$), depression ($r = .44$), and substance use ($r = .33$), with which it shares only moderate correlation (Makini et al., 1996).

Missing Data

There were substantial amounts of missing data in the Hawaiian dataset. 6.4% of all values were missing for the 1992/1993 school year, 52.2% of all values were missing for the 1993/1994 school year, and 22.53% of all values were missing for the 1994/1995 school year. The missingness was not completely at random (MCAR). We addressed this potential problem by multiply imputing missing data on all variables at the item level using SPSS version 21.0 (IBM Corp, 2012). Multiple imputation (MI) is increasingly advocated as the optimal approach for dealing with missing data (Graham, 2009; Schafer & Graham, 2002; Shrive, Stuart, Quan, & Ghali, 2006). When MI has been compared with alternative methods of handling incomplete data (e.g., single imputation methods, complete-case analyses, maximum likelihood approaches), it has been shown to generate less biased estimates that have more statistical efficiency (e.g., Crawford, Tennstedt, & McKinlay, 1995; Donders, van der Heijden, Stijnen, & Moons, 2006; Liu & Gould, 2002; Tang, Song, Belin, & Unutzer, 2005). There is also evidence indicating that MI performs well across different circumstances, such as small samples, very large multiple regressions, and when there are large amounts of missing data (Graham & Schafer, 1999).

MI works by generating plausible missing values multiple times based on the distribution of the observed data. Random components are incorporated into these estimated values to reflect their uncertainty. This procedure creates a set of "complete" data sets with no missing values. Analyses are then run separately on each data set, and the results are pooled across data sets using multiple imputation combining rules (Enders, 2010; Graham, 2009). The purpose of MI is not to obtain the individual values themselves but to estimate unbiased parameter estimates of the data set as a whole (Graham, 2009). We followed recommendations to match the number of imputations to the fraction of missing information because progressively larger numbers of imputed data sets are needed to maximize power in subsequent significance testing (Bodner, 2008; Graham, Olchowski, & Gilreath, 2007; White et al., 2011).

Statistical Analysis

Analyses were conducted using SPSS version 21.0 (IBM Corp, 2012) and R (R Development Core Team, 2009). Confirmatory Factor Analysis (CFA) was performed using the R lavaan package, version 0.5–18 (Rosseel, 2012). Three CFA models were tested using full information maximum likelihood (ML) estimation. Model 1 was the standard two factor model involving separate negatively worded items ("anxiety present") and positively worded items ("anxiety absent"), which were allowed to correlate. Model 2 was a single factor model with all items loading on a single factor. Model 3 featured a single substantive anxiety/calmness factor, but all positively worded items were allowed to cross-load onto a second methodological artifact factor which takes into account additional residual intercorrelation between positively worded items. The three

CFA models were estimated separately for the state and trait subscales and the two samples to ensure that findings were not specific to a particular form of the STAI or sample.

Acceptable fit was operationalized as Root Mean Squared Error of Approximation (RMSEA) ≤ .08, Comparative Fit Index (CFI) ≥ .90, Tucker Lewis Index (TLI) ≥ .90. Good fit was operationalized as RMSEA ≤ .06, CFI ≥ .95, and TLI ≥ .95 (Hu & Bentler, 1999). Competing models were compared using (a) Akaike's Information Criterion (AIC), which tests the relative fit of competing models after adjusting for parsimony (lower AICs indicate less information loss and thus a superior model), (b) CFI, using a .002 cutoff (Meade, Johnson, & Braddy, 2008), and (c) Bayesian Information Criterion (BIC), where lower BIC statistics suggest better fit. Although the AIC and BIC share the same goodness-of-fit term, the penalty term of BIC is potentially much more stringent than the penalty term of AIC so BIC tends to choose fitted models that are more parsimonious than those favored by AIC. The AIC, CFI, and BIC model comparison indices have the advantage of being less compromised by large sample sizes when compared with the chi-squared and chi-squared difference statistics (see Cheung & Rensvold, 2002; Meade et al., 2008).

As STAI data are ordinal, we conducted a robustness check of our CFA results by replicating them using mean- and variance-adjusted weighted least squares (WLSMV) estimation. WLSMV estimation has been found to result in unbiased parameter and standard error estimates, and acceptable type-I error rates for structural equation modeling with (skewed) ordinal variables (Flora & Curran, 2004). Competing models were compared using change in model fit according to CFI and RMSEA (Chen et al., 2008).

Hierarchical ordinary least squares (OLS) regressions were used to explore linear and nonlinear relationships between STAI trait scores, treated as a single factor (all items summed to produce a total trait score), and outcome variables. In each analysis, Step 1 involved fitting a model whereby STAI trait scores had a linear relationship with each outcome variable measured at the same time (1992/1993 school year), or measured at follow-up 1 or 2 years later (1993/1994, 1994/1995 school years), while controlling for scores on the outcome variable at baseline (hence it was the change in outcome that we were predicting). Steps 2 and 3 tested whether adding a nonlinear term (squared and cubed STAI trait total scores) made a significant improvement to the amount of variance explained. Improvement in model fit was based on ΔR^2. Statistically significant deviations from linearity were graphed to visually display relationships, using unstandardized regression coefficients. This also clarified whether nonlinearity was substantive. The state score, by its very nature, was not expected to reliably predict changes in outcome variables over time and so did not feature in these analyses.

Results

Comparison of CFA Models

Table 1 shows that across state and trait items and both samples, the two factor model (Model 1) demonstrated an improvement in fit over the single factor model (Model 2). This replicates previous findings. However, Model 3 subsequently outperformed the traditional two factor model, suggesting that when shared method bias among positively worded items is controlled for, a single factor underlies the STAI items. The AIC and BIC statistics, which account for model complexity, and the change in model fit according to CFI and RMSEA (Chen et al., 2008), all indicated superiority of Model 3 as hypothesized. Overall fit was acceptable for Model 3 in all instances, with the exception of the state items in the Hawai‘i sample, where fit indices fell slightly below our criteria. Nonetheless the adjusted model demonstrated better fit than the traditional two factor model and was therefore favored.

Table 1
Comparison of Three Maximum Likelihood CFA Models in Two Independent Samples

Model	Model fit						
	χ^2	*df*	AIC	BIC	TLI	CFI	RMSEA
British sample, trait items ($N = 1{,}824$)							
1. Two factor	1521.088*	169	105393.044	105618.904	.938	.945	.066
2. Single factor	2297.308*	170	106167.264	106387.615	.903	.913	.083
3. Single factor, method variance factor	1399.848*	163	105283.803	105542.716	.941	.950	.064
British sample, state items ($N = 1{,}824$)[a]							
1. Two factor	1778.852*	169	77214.564	77429.490	.928	.936	.083
2. Single factor	2773.269*	170	78206.982	78416.665	.884	.896	.105
3. Single factor, method variance factor	1464.346*	160	76918.058	77180.162	.938	.948	.076
Hawai‘i sample, trait items ($N = 4{,}138$)[b]							
1. Two factor	1882.202*	169	150252.538	150503.738	.926	.935	.055
2. Single factor	8382.153*	170	156750.489	156995.562	.649	.686	.119
3. Single factor, method variance factor	1773.737*	163	150156.073	150444.033	.923	.938	.054
Hawai‘i sample, state items ($N = 4{,}138$)[b]							
1. Two factor	4987.366*	169	160309.299	160562.289	.842	.859	.090
2. Single factor	11325.305*	170	166645.238	166892.060	.636	.674	.136
3. Single factor, method variance factor	4744.901*	160	160084.834	160393.357	.841	.866	.090

Note. AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; CFI = Comparative Fit Index; RMSEA = Root Mean Squared Error of Approximation; TLI = Tucker Lewis Index.
[a] STAI at six weeks post natal. [b] STAI completed during 1992/1993 school year. Analyses are reported to three decimal places for clarity.
* $p < .001$.

Our robustness check employing WLSMV estimation corroborated these results (see supplementary material). Again, the two factor model demonstrated an improvement in fit over the single factor model in all instances. Model 3 outperformed the traditional two factor model in the British sample and performed similarly in the Hawai‘i sample. Overall fit was acceptable for Model 3 in the majority of instances. Taken together, the CFA results using ML and WLSMV estimation support the superiority of Model 3 (a calmness–anxiety continuum) over the traditional two factor model (separate "anxiety present" and "anxiety absent" factors), suggesting that when shared method bias among positively worded items is controlled for, a single factor underlies STAI state and trait scales.

Exploration of Linear and Nonlinear Relationships With Outcome Variables

A series of regression analyses were conducted in the Hawai‘i sample to explore the form of the relationship between STAI trait scores and outcome variables over time (see Table 2). Three of the nine regression models showed statistically significant nonlinear relationships for Step 2. However, the squared term accounted for very little additional variation above and beyond the linear main effect (2.4%, 1.6%, 0.1%). Two of the nine regression models showed statistically significant nonlinear relationships for Step 3. However, again, the cubed term accounted for very little additional variation (0.2%, 0.2%). Thus, in all cases the nonlinear term failed to make any substantive improvement to the original linear model and these results provide only very weak evidence of a nonlinear relationship.

Potential nonlinearity was further explored by graphing statistically significant nonlinear relationships (see Figure 1). The graphs reveal only subtle variation away from perfect linearity. Given the large proportion of missing data, we conducted a robustness check of our regression analyses using complete cases (see supplementary material). These results were almost identical (in terms of ΔR^2 values), again finding that baseline levels on the calmness–anxiety continuum have a near linear relationship with changes in depression, aggression, and substance misuse over time (regression coefficients are larger than those in Table 2 as multiple imputation estimates relationships more conservatively).

Overall, the regression analyses and graphs provide evidence that the calmness–anxiety continuum (STAI trait total scores) has a near linear relationship with outcome variables over time. Any nonlinearity appears to be of statistical but not practical or clinical significance. These results support our prediction that moving from anxiety to calmness on the STAI provides an equal decrease in risk for several other psychological problems, irrespective of position on the calmness–anxiety continuum.

Discussion

Our results replicate those of Vautier and Pohl (2009). Like them, we demonstrated that state and trait anxiety, as measured by the STAI, can be understood as continua that range from high calmness to high anxiety. Our analyses were underpinned by Joseph and Wood's (2010) hypothesis that STAI "anxiety absent" items (e.g., "I am cool, calm and collected" [trait anxiety]; "I feel calm" [state anxiety]) assess the presence of calmness, rather than the mere absence of anxiety problems. We provided the first evidence to corroborate this hypothesis. These results have clear implications for the structure and definition of anxiety as they go against the view that anxiety ranges from zero to intense.

We established an anxiety-calmness continuum using the English-language version of the STAI in mixed samples of adults, adolescents, and different ethnic groups from opposite sides of the globe (Hawai‘i and the U.K.) and replicated our results using both ML and WLSMV estimation. It was important to clarify that anxiety-calmness continua existed across diverse circumstances because the mechanisms underlying anxiety may differ across groups (e.g., Field & Lester, 2010; Kirmayer, Young, & Hayton, 1995; Manson, 1996), which could lead to misleading artifacts of noninvariant measurement (e.g., item content or wording that is biased against a given group). The use of a large dataset that is representative of adolescents residing in Hawaii minimized the likelihood of systematic sampling bias, which could have been introduced had we used a purely community or clinical sample.

Previous factor analytic evidence for separate "anxiety present" and "anxiety absent" factors likely arose because there is additional common variance between positively worded items that is unrelated to the underlying latent variable. Based on item content, the STAI state and trait continua can be reconceptualized as calmness–anxiety continua (Joseph & Wood, 2010; Vautier & Pohl, 2009). This conceptualization is more intuitive than the "anxiety present" and "anxiety absent" conceptualization. We note that it is still appropriate to continue to reverse score positively worded STAI items to produce total state and trait scores. However, in light of the present findings, STAI users are advised to interpret total state and trait scores as an indication of anxiety problem severity that is based on a combination of the presence/absence of anxiety problems and the presence/absence of calmness, understanding that as anxiety problems increase, calmness decreases (and vice versa).

As the STAI shows high convergent validity with other leading measures of anxiety problems (Barnes et al., 2002; Knight et al., 1983; Spielberger, 1989), a calmness–anxiety continuum may be apparent in other scales that contain factors that consist of entirely positively and negatively worded items. Scales which measure anxiety or calmness, but which do not contain a mixture of positively and negatively worded items, are presumably measuring one half of the continuum. Evidence of a calmness–anxiety continuum suggests that existing research into anxiety problems will have relevance for the field of calmness and relaxation research, and vice versa, and that studying anxiety or calmness separately may be unnecessarily duplicating research effort (Joseph & Wood, 2010).

This is the first study to characterize the form of the relationship between the calmness–anxiety continuum and other psychiatric variables. These analyses make the clinical importance of this topic more apparent and explicit than merely examining whether a continuum exists as Vautier and Pohl (2009) did. Our results demonstrate that baseline levels on the calmness–anxiety continuum have a near linear relationship with outcome variables measured at the same time and one and two years later. Thus, there is no intrinsic way to demarcate problematic degrees of anxiety (or beneficial degrees of calmness) based on how anxiety or calmness are related to other psychiatric variables. That these results were

Table 2
Results of Regression Analyses Comparing Linear and Nonlinear Effects of STAI Upon Change in Outcome

Step	Variable	*B*	*SE B*	β	ΔR^2
	Depression as outcome				
1992/1993 school year (*N* = 4,069)					
1	Constant	−1.569	.277		
	Total STAI Trait score	.788	.011	.744***	.554***
2	Constant	4.235	.470		
	Total STAI Trait score	.209	.040	.198***	
	Total STAI Trait score squared	.012	.001	.567***	.024***
3	Constant	6.740	.719		
	Total STAI Trait score	−.215	.101	.095*	
	Total STAI Trait score squared	.031	.004	.201***	
	Total STAI Trait score cubed	.000	.000	.116***	.002***
1993/1994 school year (*N* = 4,101)					
1	Constant	12.715	.398		
	1992/1993 CES-D total score	.199	.022	.246***	
	Total STAI Trait score	.159	.024	.164***	.149***
2	Constant	13.141	.668		
	1992/1993 CES-D total score	.222	.023	.242***	
	Total STAI Trait score	.119	.056	.122*	
	Total STAI Trait score squared	.001	.001	.046	.000
3	Constant	12.702	1.011		
	1992/1993 CES-D total score	.223	.023	.025***	
	Total STAI Trait score	.192	.139	.143	
	Total STAI Trait score squared	−.002	.006	.304	
	Total STAI Trait score cubed	.000	.000	.176	.000
1994/1995 school year (*N* = 4,101)					
1	Constant	19.015	.568		
	1992/1993 CES-D total score	.153	.027	.197***	
	Total STAI Trait score	.083	.027	.100***	.081***
2	Constant	18.687	.774		
	1992/1993 CES-D total score	.156	.027	.201***	
	Total STAI Trait score	.113	.058	.137*	
	Total STAI Trait score squared	−.001	.001	−.042	.000
3	Constant	18.354	1.069		
	1992/1993 CES-D total score	.157	.027	.202***	
	Total STAI Trait score	.169	.135	.205	
	Total STAI Trait score squared	−.003	.006	−.194	
	Total STAI Trait score cubed	.000	.000	.089	.000
	Aggression as outcome				
1992/1993 school year (*N* = 4,069)					
1	Constant	−.053	.169		
	Total STAI Trait score	.260	.007	.515***	.266***
2	Constant	2.191	.292		
	Total STAI Trait score	.036	.025	.072	
	Total STAI Trait score squared	.005	.000	.461***	.016***
3	Constant	2.747	.447		
	Total STAI Trait score	−.058	.063	−.115	
	Total STAI Trait score squared	.009	.003	.881***	
	Total STAI Trait score cubed	.001	.000	−.247	.001
1993/1994 school year (*N* = 4,101)					
1	Constant	4.903	.192		
	1992/1993 BADS total score	.341	.017	.367***	
	Total STAI Trait score	.042	.009	.089***	.175***
2	Constant	4.880	.321		
	1992/1993 BADS total score	.341	.018	.367***	
	Total STAI Trait score	.044	.027	.094	
	Total STAI Trait score squared	.000	.001	−.005	.000
3	Constant	4.951	.484		
	1992/1993 BADS total score	.341	.018	.367***	
	Total STAI Trait score	.032	.067	.068	
	Total STAI Trait score squared	.000	.003	.053	
	Total STAI Trait score cubed	.000	.000	−.034	.000

(*table continues*)

Table 2 (*continued*)

Step	Variable	*B*	*SE B*	β	ΔR^2
1994/1995 school year (N = 4,101)					
1	Constant	7.743	.256		
	1992/1993 BADS total score	.244	.023	.302***	
	Total STAI Trait score	.021	.010	.051*	.311*
2	Constant	7.416	.342		
	1992/1993 BADS total score	.247	.023	.306***	
	Total STAI Trait score	.053	.025	.129*	
	Total STAI Trait score squared	−.001	.001	−.083	.000
3	Constant	7.452	.484		
	1992/1993 BADS total score	.247	.023	.306***	
	Total STAI Trait score	.047	.063	.114	
	Total STAI Trait score squared	.000	.003	−.050	
	Total STAI Trait score cubed	.000	.000	−.019	.000
Substance misuse as outcome					
1992/1993 school year (N = 4,069)					
1	Constant	.005	.053		
	Total STAI Trait score	.044	.002	.308***	.095***
2	Constant	.170	.092		
	Total STAI Trait score	.027	.008	.192***	
	Total STAI Trait score squared	.000	.000	.120*	.001*
3	Constant	.469	.141		
	Total STAI Trait score	−.024	.020	−.166	
	Total STAI Trait score squared	.003	.001	.928***	
	Total STAI Trait score cubed	.000	.000	−.474***	.002*
1993/1994 school year (N = 4,101)					
1	Constant	1.111	.053		
	1992/1993 SASSI-A total score	.403	.016	.442***	
	Total STAI Trait score	.005	.002	.036*	.204***
2	Constant	1.125	.085		
	1992/1993 SASSI-A total score	.403	.016	.442***	
	Total STAI Trait score	.003	.007	.025	
	Total STAI Trait score squared	.000	.000	.011	.000
3	Constant	.983	.130		
	1992/1993 SASSI-A total score	.404	.016	.443***	
	Total STAI Trait score	.027	.018	.211	
	Total STAI Trait score squared	−.001	.001	−.407	
	Total STAI Trait score cubed	.000	.000	.246	.000
1994/1995 school year (N = 4,101)					
1	Constant	1.714	.067		
	1992/1993 SASSI-A total score	.222	.020	.298***	
	Total STAI Trait score	.005	.002	.051*	.097***
2	Constant	1.650	.094		
	1992/1993 SASSI-A total score	.222	.020	.299***	
	Total STAI Trait score	.012	.007	.112	
	Total STAI Trait score squared	.000	.000	−.064	.000
3	Constant	1.532	.133		
	1992/1993 SASSI-A total score	.223	.020	.300***	
	Total STAI Trait score	.032	.018	.301	
	Total STAI Trait score squared	−.001	.001	−.489	
	Total STAI Trait score cubed	.000	.000	.249	.001

Note. STAI = State Trait Anxiety Inventory; CES-D = Centre for Epidemiological Studies-Depression; BADS = Braver Aggressiveness Dimension Scale; SASSI-A = Substance Abuse Subtle Screening Inventory—Adolescent version; STAI Trait scale completed during 1992/1993 school year. Analyses are reported to three decimal places for clarity.
* $p < .05$. *** $p < .001$.

apparent over time suggests that moving along the continuum toward high calmness provides continuous and long-term protection against experiencing other psychological problems.

The present study provides evidence to substantiate calls by professional bodies (e.g., The British Psychological Society, 2010), the Positive Clinical Psychology movement (e.g., Wood & Tarrier, 2010, as clarified in Johnson & Wood, in press; Wood & Johnson, 2016), and many clinicians, for clinical services to adopt a broader focus that jointly involves reducing distress *and* increasing well-being. Our results point to the usefulness of early intervention and prevention (when people begin to move away from high calmness) and instilling resilience (by providing interventions to move people toward high calmness). Fostering high levels of calmness would mean that individuals have further to go before

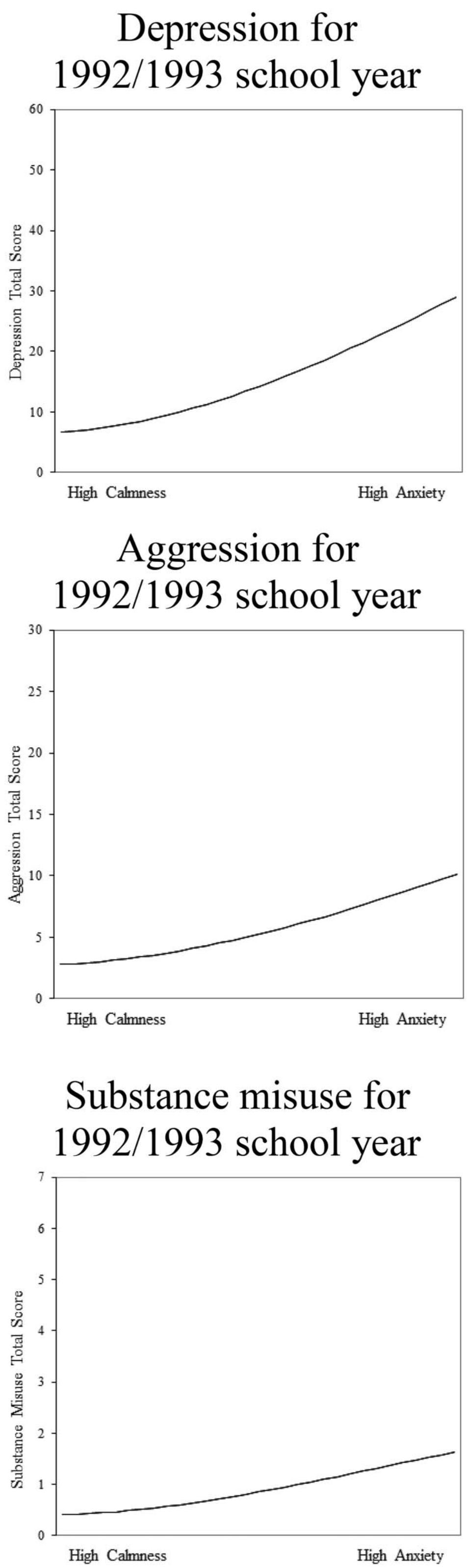

Figure 1. Line graphs plotting unstandardized nonlinear regression lines for statistically significant ΔR^2 values. Total STAI trait scores predict outcome variables at different time points; STAI trait scale completed during 1992/1993 school year.

they reach high levels of anxiety. Psychiatric and psychological interventions that are grounded in a continuum conceptualization would logically be stopped when an individual reports high calmness. Patients, service commissioners and others may of course want interventions to stop at the point of problem absence. However, the present results provide an evidence base to inform a collaborative discussion around when to stop treatment and the advantages and disadvantages of doing so. This process may already be happening when clinicians construct relapse prevention plans and offer booster sessions with people who have finished treatment.

We are hopeful that our findings could help support a case for publically funded clinical services to accept the promotion of well-being into their remit. That well-being interventions often fall outside the focus of publically funded clinical services means that efforts to help the public address this need (e.g., through self-help books, self-development courses, and other “interventions”) are often offered by unaccredited and untrained individuals. We find this concerning, as we strongly believe that clinical interventions should be targeted, theory-driven, evidence-based, and provided by suitably qualified, ethically practicing professionals.

Our results could be extended in a number of ways. Because cost considerations have a substantial impact on service delivery and often outweigh theoretical or moral arguments, the current results need to be accompanied by a comprehensive economic cost value analysis which tests the implications of the continuum conceptualization and when it is most cost-effective to stop interventions. It is obviously less cost-effective in the short-term to stop interventions at the point of well-being rather than mere problem absence, or to start interventions when people begin to move away from high calmness but before a severe psychological problem becomes manifest. However, this approach may prove to be the most cost-effective solution overall if it provides long-term protection from other problems, especially among high risk groups. Research is also needed to characterize the form of the relationship between the calmness–anxiety continuum and other psychiatric variables in relation to all age groups who complete the STAI.

We also note that the current findings were limited in not accounting for random and systematic measurement error (Barrett & Russell, 1998; Green et al., 1993). Future research investigating the calmness–anxiety continuum could account for intraindividual mood variation and measurement error by taking repeated (e.g., daily) continuous measures of mood using different response formats.

References

Achenbach, T. M. (1991). *Manual for the Youth Self-Report and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Amtmann, D., Kim, J., Chung, H., Bamer, A. M., Askew, R. L., Wu, S., . . . Johnson, K. L. (2014). Comparing CESD-10, PHQ-9, and PROMIS depression instruments in individuals with multiple sclerosis. *Rehabilitation Psychology, 59,* 220–229. http://dx.doi.org/10.1037/a0035919

Andrade, N. N., Hishinuma, E. S., McDermott, J. F., Jr., Johnson, R. C., Goebert, D. A., Makini, G. K., Jr., . . . Waldron, J. A. (2006). The National Center on Indigenous Hawaiian Behavioral Health study of prevalence of psychiatric disorders in native Hawaiian adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 45,* 26–36. http://dx.doi.org/10.1097/01.chi.0000184933.71917.f4

Barnes, L. L. B., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the Spielberger State–Trait Anxiety Inventory. *Educational and Psychological Measurement, 62,* 603–618. http://dx.doi.org/10.1177/0013164402062004005

Barrett, L. F., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology, 74,* 967–984. http://dx.doi.org/10.1037/0022-3514.74.4.967

Bernstein, I. H., & Eveland, D. C. (1982). State vs. trait anxiety: A case study in confirmatory factor analysis. *Personality and Individual Differences, 3,* 361–372. http://dx.doi.org/10.1016/0191-8869(82)90002-2

Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural Equation Modeling, 15,* 651–675. http://dx.doi.org/10.1080/10705510802339072

Boeke, S., Duivenvoorden, H., & Bonke, B. (1984). Agreement for surgical patients on two situations for the trait anxiety inventory. *Psychological Reports, 54,* 278. http://dx.doi.org/10.2466/pr0.1984.54.1.278

Braver, S. L., Fogas, B., Sandler, I., & Volchik, S. (1986, May). *Development of a self-report aggressiveness scale for children: The Braver Aggressiveness Dimension Scale (BADS).* Paper presented at the annual meeting of the Western Psychological Association, Seattle, WA.

Broadbent, D. E. (1965). A reformulation of the Yerkes-Dodson law. *British Journal of Mathematical and Statistical Psychology, 18,* 145–157. http://dx.doi.org/10.1111/j.2044-8317.1965.tb00338.x

Canals, J., Martí-Henneberg, C., Fernández-Ballart, J., Clivillé, R., & Domènech, E. (1992). Scores on the State-Trait Anxiety Inventory for children in a longitudinal study of pubertal Spanish youth. *Psychological Reports, 71,* 503–512. http://dx.doi.org/10.2466/pr0.1992.71.2.503

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior.* New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139174794

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research, 36,* 462–494. http://dx.doi.org/10.1177/0049124108314720

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5

Crawford, S. L., Tennstedt, S. L., & McKinlay, J. B. (1995). A comparison of analytic methods for non-random missingness of outcome data. *Journal of Clinical Epidemiology, 48,* 209–219. http://dx.doi.org/10.1016/0895-4356(94)00124-9

Donders, A. R. T., van der Heijden, G. J. M. G., Stijnen, T., & Moons, K. G. M. (2006). Review: A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology, 59,* 1087–1091. http://dx.doi.org/10.1016/j.jclinepi.2006.01.014

Enders, C. K. (2010). *Applied missing data analysis.* New York, NY: Guilford Press.

Ensel, W. M. (1986). Measuring depression: The CES-D Scale. In N. Lin, A. Dean, & W. M. Ensel (Eds.), *Social support, life events, and depression* (pp. 51–70). New York, NY: Academic Press

Field, A. P., & Lester, K. J. (2010). Is there room for 'development' in developmental models of information processing biases to threat in children and adolescents? *Clinical Child and Family Psychology Review, 13,* 315–332. http://dx.doi.org/10.1007/s10567-010-0078-8

Flett, G. L., Vredenburg, K., & Krames, L. (1997). The continuity of depression in clinical and nonclinical samples. *Psychological Bulletin, 121,* 395–416. http://dx.doi.org/10.1037/0033-2909.121.3.395

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9,* 466–491. http://dx.doi.org/10.1037/1082-989X.9.4.466

Garrison, C. Z., Addy, C. L., Jackson, K. L., McKeown, R. E., & Waller, J. L. (1991). The CES-D as a screen for depression and other psychiatric disorders in adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 30,* 636–641. http://dx.doi.org/10.1097/00004583-199107000-00017

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549–576. http://dx.doi.org/10.1146/annurev.psych.58.110405.085530

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8,* 206–213. http://dx.doi.org/10.1007/s11121-007-0070-9

Graham, J. W., & Schafer, J. L. (1999). On the performance of multiple imputation for multivariate data with small sample size. In R. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 1–29). Thousand Oaks, CA: Sage.

Green, D. P., Goldman, S. L., & Salovey, P. (1993). Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology, 64,* 1029–1041. http://dx.doi.org/10.1037/0022-3514.64.6.1029

Green, J. M., Statham, H. and Snowdon, C., Cambridge Prenatal Screening Study: Main Study, 1990–1991 [computer file]. Colchester, Essex: UK Data Archive [distributor], February 1996. SN: 3462. http://dx.doi.org/10.5255/UKDA-SN-3462-1

Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T., Goebert, D. A., Yuen, N. Y. C., Makini, G. K., Jr., . . . Carlton, B. S. (2001). Prediction of anxiety disorders using the state-trait anxiety inventory for multiethnic adolescents. *Journal of Anxiety Disorders, 15,* 511–533. http://dx.doi.org/10.1016/S0887-6185(01)00079-2

Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T., & Nahulu, L. B. (2000). Differences in State-Trait Anxiety Inventory scores for ethnically diverse adolescents in Hawaii. *Cultural Diversity and Ethnic Minority Psychology, 6,* 73–83. http://dx.doi.org/10.1037/1099-9809.6.1.73

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

IBM Corp. (2012). *IBM SPSS Statistics for Windows, Version 21.0.* Armonk, NY: Author.

Johnson, J., & Wood, A. M. (in press). Integrating positive and clinical psychology: Viewing human functioning as continua from positive to negative can benefit clinical assessment, interventions and understandings of resilience. *Cognitive Therapy and Research.*

Joseph, S., & Wood, A. (2010). Assessment of positive functioning in clinical psychology: Theoretical and practical issues. *Clinical Psychology Review, 30,* 830–838. http://dx.doi.org/10.1016/j.cpr.2010.01.002

Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry, 60,* 929–937. http://dx.doi.org/10.1001/archpsyc.60.9.929

Kirmayer, L. J., Young, A., & Hayton, B. C. (1995). The cultural context of anxiety disorders. *Psychiatric Clinics of North America, 18,* 503–521.

Knight, R. G., Waal-Manning, H. J., & Spears, G. F. (1983). Some norms and reliability data for the State—Trait Anxiety Inventory and the Zung Self-Rating Depression scale. *British Journal of Clinical Psychology, 22,* 245–249. http://dx.doi.org/10.1111/j.2044-8260.1983.tb00610.x

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin, 139,* 917–940. http://dx.doi.org/10.1037/a0030811

Liu, G., & Gould, A. L. (2002). Comparison of alternative strategies for analysis of longitudinal trials with dropouts. *Journal of Biopharmaceutical Statistics, 12,* 207–226. http://dx.doi.org/10.1081/BIP-120015744

Lonner, W. J., & Ibrahim, F. A. (1989). Assessment in cross-cultural counseling. In P. B. Pedersen, J. G. Draguns, W. J. Lonner, & J. E. Trimble (Eds.), *Counseling across cultures* (3rd ed., pp. 299–333). Honolulu, HI: University of Hawaii Press.

Makini, G. K., Jr., Andrade, N. N., Nahulu, L. B., Yuen, N., Yate, A., McDermott, J. F., Jr., . . . Waldron, J. A. (1996). Psychiatric symptoms of Hawaiian adolescents. *Cultural Diversity and Mental Health, 2,* 183–191. http://dx.doi.org/10.1037/1099-9809.2.3.183

Manson, S. M. (1996). Culture and *DSM–IV*: Implications for the diagnosis of mood and anxiety disorders. In J. E. Mezzich, A. Kleinman, H. Fabrega, Jr., & D. L. Parron (Eds.), *Culture and psychiatric diagnosis: A DSM–IV perspective* (pp. 99–113). Washington, DC: American Psychiatric Association.

Markon, K. E. (2010). How things fall apart: Understanding the nature of internalizing through its relationship with impairment. *Journal of Abnormal Psychology, 119,* 447–458. http://dx.doi.org/10.1037/a0019707

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93,* 568–592. http://dx.doi.org/10.1037/0021-9010.93.3.568

Miller, G. (1990). *The SASSI adolescent manual*. Spencer, IN: The Spencer Evening World.

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49,* 377–412. http://dx.doi.org/10.1146/annurev.psych.49.1.377

Mook, J., Van der Ploeg, H. M., & Kleijn, W. C. (1992). Symptom-positive and symptom-negative items in the State-Trait Anxiety Inventory: A comparison and replication. *Anxiety, Stress, & Coping, 5,* 113–123. http://dx.doi.org/10.1080/10615809208250491

Nishimura, S. T., Hishinuma, E. S., Miyamoto, R. H., Goebert, D. A., Johnson, R. C., Yuen, N. Y. C., & Andrade, N. N. (2001). Prediction of DISC substance abuse and dependency for ethnically diverse adolescents. *Journal of Substance Abuse, 13,* 597–607. http://dx.doi.org/10.1016/S0899-3289(01)00088-8

Piazza, N. J. (1996). Dual diagnosis and adolescent psychiatric inpatients. *Substance Use & Misuse, 31,* 215–223. http://dx.doi.org/10.3109/10826089609045809

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385–401. http://dx.doi.org/10.1177/014662167700100306

R Development Core Team. (2009). *R project for statistical computing*. Vienna, Austria: Author.

Risberg, R. A., Stevens, M. J., & Graybill, D. F. (1995). Validating the adolescent form of the Substance Abuse Subtle Screening Inventory. *Journal of Child & Adolescent Substance Abuse, 4,* 25–42. http://dx.doi.org/10.1300/J029v04n04_02

Roberts, R. E. (1980). Reliability of the CES-D Scale in different ethnic contexts. *Psychiatry Research, 2,* 125–134. http://dx.doi.org/10.1016/0165-1781(80)90069-4

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48,* 1–36. http://dx.doi.org/10.18637/jss.v048.i02

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. http://dx.doi.org/10.1037/h0077714

Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin, 125,* 3–30. http://dx.doi.org/10.1037/0033-2909.125.1.3

Salehi, B., Cordero, M. I., & Sandi, C. (2010). Learning under stress: The inverted-U-shape function revisited. *Learning & Memory (Cold Spring Harbor, N. Y.), 17,* 522–530. http://dx.doi.org/10.1101/lm.1914110

Santor, D. A., Gregus, M., & Welch, A. (2006). Eight decades of measurement in depression. *Measurement: Interdisciplinary Research and Perspectives, 4,* 135–155.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177. http://dx.doi.org/10.1037/1082-989X.7.2.147

Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement, 9,* 367–373. http://dx.doi.org/10.1177/014662168500900405

Shean, G., & Baldwin, G. (2008). Sensitivity and specificity of depression questionnaires in a college-age sample. *The Journal of Genetic Psychology: Research and Theory on Human Development, 169,* 281–292. http://dx.doi.org/10.3200/GNTP.169.3.281-292

Shrive, F. M., Stuart, H., Quan, H., & Ghali, W. A. (2006). Dealing with missing data in a multi-question depression scale: A comparison of imputation methods. *BMC Medical Research Methodology, 6,* 57–67. http://dx.doi.org/10.1186/1471-2288-6-57

Spielberger, C. D. (1989). *State–Trait Anxiety Inventory: A comprehensive bibliography* (2nd ed.). Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., & Diaz-Guerrero, R. (Eds.). (1986). *Cross-cultural anxiety* (Vol. 3). Washington, DC: Hemisphere/Wiley.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *State–Trait Anxiety Inventory for adults (Form X)*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Vagg, P. R., Barker, L. R., Donham, G. W., & Westberry, L. G. (1980). The factor structure of the State–Trait Anxiety Inventory. In I. G. Sarason & C. D. Spielberger (Eds.), *Stress and anxiety* (Vol. 7, pp. 95–109). New York, NY: Hemisphere/Wiley.

Suzuki, T., Tsukamoto, K., & Abe, K. (2000). Characteristics factor structures of the Japanese version of the State-Trait Anxiety Inventory: Coexistence of positive-negative and state-trait factor structures. *Journal of Personality Assessment, 74,* 447–458. http://dx.doi.org/10.1207/S15327752JPA7403_8

Tang, L., Song, J., Belin, T. R., & Unützer, J. (2005). A comparison of imputation methods in a longitudinal randomized clinical trial. *Statistics in Medicine, 24,* 2111–2128. http://dx.doi.org/10.1002/sim.2099

The British Psychological Society. (2010). *Division of Clinical Psychology: The core purpose and philosophy of the profession*. Leicester, UK: The British Psychological Society.

Vagg, P. R., Spielberger, C. D., & O'Hearn, T. P. J. R., Jr. (1980). Is the State–Trait anxiety inventory multidimensional? *Personality and Individual Differences, 1,* 207–214. http://dx.doi.org/10.1016/0191-8869(80)90052-5

Vautier, S., & Pohl, S. (2009). Do balanced scales assess bipolar constructs? The case of the STAI scales. *Psychological Assessment, 21,* 187–193. http://dx.doi.org/10.1037/a0015312

Vigneau, F., & Cormier, S. (2008). The factor structure of the State-Trait Anxiety Inventory: An alternative view. *Journal of Personality Assessment, 90,* 280–285. http://dx.doi.org/10.1080/00223890701885027

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine, 30,* 377–399. http://dx.doi.org/10.1002/sim.4067

Wood, A. M., & Johnson, J. (Eds.). (2016). *The Wiley handbook of positive clinical psychology: An integrative approach to studying and improving well-being*. Chichester, UK: Wiley Ltd. http://dx.doi.org/10.1002/9781118468197

Wood, A. M., & Tarrier, N. (2010). Positive Clinical Psychology: A new vision and strategy for integrated research and practice. *Clinical Psychology Review, 30,* 819–829. http://dx.doi.org/10.1016/j.cpr.2010.06.003

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment, 28,* 186–194. http://dx.doi.org/10.1007/s10862-005-9004-7

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology & Psychology, 18,* 459–482.

Yik, M., Russell, J. A., & Steiger, J. H. (2011). A 12-point circumplex structure of core affect. *Emotion, 11,* 705–731. http://dx.doi.org/10.1037/a0023980

Received February 4, 2016
Revision received October 19, 2016
Accepted November 3, 2016 ■