Psychological Assessment
2010, Vol. 22, No. 4, 878–892

1040-3590/10/$12.00 DOI: 10.1037/a0020644

Self-Discrepancy: Comparisons of the Psychometric Properties of Three Instruments

Neill Watson, Brandon C. Bryan, and Todd M. Thrash
College of William and Mary

In 2 studies, the psychometric properties of 3 methods for measuring real–ideal and real–ought self-discrepancies were compared: the idiographic Self-Concept Questionnaire–Personal Constructs, the nonidiographic Self-Concept Questionnaire–Conventional Constructs, and the content-free Abstract Measures. In the 1st study, 125 students at a university clinic completed the 3 instruments and measures of anxiety and depression before individual therapy. In the 2nd study, 278 undergraduates completed the 3 instruments at 2 time points 4 weeks apart and completed multiple measures of anxiety and depression at the 2nd time point. Internal consistency alphas were consistently strong for the personal construct measures (.90 to .92) and moderate to strong for the conventional construct measures (.82 to .90). Test–retest reliability coefficients were above .70 for the personal construct and conventional construct measures, but the coefficients for the latter were inflated by the stability of their error terms. The 2 discrepancies were found to be factorially distinct even though they were highly correlated. Convergent and discriminant evidence of validity was found in both studies for all measures except the abstract real–ought discrepancy. Convergence was as strong or stronger for the personal construct measures in comparison to the other measures. Test-criterion evidence of validity, with multiple measures of anxiety and depression as criteria, was found in both studies for all measures except for the abstract real–ought discrepancy in relation to anxiety. Overall, the findings support the idiographic personal construct instrument most strongly for clinical assessment and for clinical, translational, and personality research.

Keywords: self-discrepancy, psychometric properties, anxiety, depression, personal constructs

The relationship of self-discrepancy to psychological distress has been an area of research since Rogers introduced the discrepancy between the real self (the self as one sees oneself) and the ideal self (the self as one would like to be in one's own eyes) as a psychotherapy outcome measure (Rogers & Dymond, 1954). Higgins later introduced the discrepancy between the real self and the ought self (the self as one believes others think one ought to or should be) in personality research, theorizing that the two discrepancies are distinct constructs (Higgins, 1987; Higgins, Klein, & Strauman, 1985).[1] Researchers using Higgins's Selves Questionnaire, which measures both discrepancies, have questioned whether the constructs are psychometrically distinct (Gramzow, Sedikides, Panter, & Insko, 2000; Phillips & Silvia, 2005; Tangney, Niedenthal, Covert, & Barlow, 1998) and have criticized the rationale and difficulty of its scoring (Francis, Boldero, & Sambell, 2006; Scott & O'Hara, 1993; Tangney et al., 1998). The purpose of the present study was to evaluate and compare, in clinical and nonclinical samples, the convergent and discriminant evidence of validity and other psychometric properties of three recently developed instruments that each measure the two discrepancies: an idiographic personal construct questionnaire, a nonidiographic conventional construct questionnaire, and a content-free abstract instrument.

Overview of Theory and Past Research

In his phenomenological theory of personality, James (1890/1981) identified self-discrepancy in the structure of self-concept, describing negative emotions associated with not meeting one's expectations of oneself. Rogers (1951, 1959) theorized that the introjection of conditions of worth from significant others leads to a high real–ideal discrepancy and that this discrepancy is a personality predisposition to emotional distress. Higgins (1987) theorized that the real–ought discrepancy is uniquely related to anxiety and the real–ideal discrepancy is uniquely related to depression.

Research to date on Rogers's (1951, 1959) and Higgins's (1987) theories suggests that the real-ideal (RI) and real-ought (RO) discrepancies are useful constructs in clinical and personality research and clinical practice. Before Higgins introduced RO along with RI in the Selves Questionnaire (Higgins, 1987; Higgins et al.,

Neill Watson, Brandon C. Bryan, and Todd M. Thrash, Department of Psychology, College of William and Mary. Brandon C. Bryan is now at the W. G. Hefner Veterans Affairs Medical Center, Salisbury, North Carolina.

We thank Jack Warner for his programming of the online questionnaires. We also thank Laura Ferrell for her assistance with the data collection and David Canose for his assistance with manuscript preparation. The self-discrepancy instruments can be viewed at http://www.wm.edu/research/watson.

Correspondence concerning this article should be addressed to Neill Watson, Department of Psychology, College of William and Mary, Williamsburg, VA 23187-8795. E-mail: npwats@wm.edu

[1] Higgins's term is *actual self* (Higgins, 1987; Higgins et al., 1985). Rogers's earlier term *real self* is used in the present study (Rogers & Dymond, 1954).

1985), researchers examined only RI. In research on the outcome of client-centered therapy using the Butler–Haigh Q-Sort to measure RI, Rogers and colleagues found a decrease in RI together with decreases in anxiety and depression (Barrett-Lennard, 1962; Rogers & Dymond, 1954). Meyer (1981), adapting the Giessen Test to measure RI, also found decreases in RI and symptoms in the outcome of client-centered therapy. More recently, studies have used the Selves Questionnaire to measure both RI and RO. In outcome research on cognitive–behavioral and interpersonal therapies for depression, Strauman et al. (2001) found decreases in RI and depressive symptoms. Furthermore, studies of clinical samples have found associations of RI and RO with both anxiety and depression (Fairbrother & Moretti, 1998; Kinderman & Bentall, 1996; Scott & O'Hara, 1993; Strauman, 1989; Weilage & Hope, 1999), and studies of nonclinical samples using clinical measures have found associations of RI and RO with anxiety and/or depression (e.g., Bruch, Rivet, & Laurenti, 2000; Strauman, 1992; Strauman & Higgins, 1988, Study 2; Tangney et al., 1998). The long-term stability of the two discrepancies over a 3-year period is evidence that they are enduring structures of self-concept, as expected of a personality predisposition (Strauman, 1996). These findings provide a basis for continuing personality and clinical research on high discrepancies as predispositions to emotional distress and for the use of discrepancy instruments in social psychology translational research (Tashiro & Mortensen, 2006) and clinical research on psychotherapy process and outcome.

Self-Discrepancy Instruments in the Present Study

Although research on self-discrepancy is promising, criticisms of the psychometric properties of the most widely used instrument, the Selves Questionnaire (Higgins, 1987; Higgins et al., 1985), have raised questions about its continued use. Short-term test–retest reliability coefficients of the RI and RO measures are consistently lower than .70 (Higgins, 1987; Moretti & Higgins, 1990; Scott & O'Hara, 1993), a benchmark recommended by Joiner, Walker, Pettit, Perez, and Cukrowicz (2005). Researchers have noted high correlations between the RI and RO discrepancy measures and questioned their discriminant validity, suggesting that only a single discrepancy exists (Gramzow et al., 2000; Phillips & Silvia, 2005; Tangney et al., 1998). This instrument also has been faulted for the rationale and difficulty of its scoring (Francis et al., 2006; Scott & O'Hara, 1993; Tangney et al., 1998).

For continuing research on self-discrepancy, Watson (2004, http://www.wm.edu/research/watson) developed three instruments that assess RI and RO: the idiographic Self-Concept Questionnaire–Personal Constructs (SCQ-PC), the nonidiographic Self-Concept Questionnaire–Conventional Constructs (SCQ-CC), and the content-free Abstract Measures (AM). Like the Selves Questionnaire, the first two of these instruments may be classified as *discrepancy-between-perceptions* instruments; the participant provides ratings of the real self and the ideal or ought self, and the researcher computes the discrepancy between them. In contrast, the third measure is a *perception-of-discrepancy* instrument; rather than rating the real self and ideal or ought self, the participant rates the magnitude of discrepancy between them. In the following sections, we describe the three instruments and the theoretical contexts of their development.

SCQ-PC

From the perspective of phenomenological theory, self-discrepancy is an experiential entity (James, 1890/1981; Rogers, 1951, 1959; Shlien, 1962; Wylie, 1974). The idiographic PC instrument was designed to correspond to the individual's experience of RI and RO. The content of the instrument was based on Kelly's (1955) theory of personal constructs, which emphasizes individual uniqueness in perceptions. The theory states that a person uses a term and its polar opposite to construe meaning. Accordingly, the PC instrument elicits an individual's bipolar constructs to provide the personality characteristics that are its idiographic content. The instrument elicits bipolar constructs that describe the real self, ideal self, and ought self and then asks how often each characteristic is true of each of the self-components. RI is computed as the average absolute difference between the real-self and ideal-self ratings across the real-self and ideal-self characteristics; RO is computed in parallel fashion with the real-self and ought-self characteristics (see the Method section). In the development of this instrument, a study using an early version of the RI measure, which used a different method for eliciting personal constructs, supported its test–retest reliability ($r = .81$ over 3 to 5 weeks) and its test-criterion relationship to neuroticism (Watson & Watts, 2001). A nonpsychometric virtue of the PC instrument is that a therapist can use its idiographic content to increase empathic communication.

SCQ-CC

The nonidiographic CC instrument uses personality characteristics from Parker and Veldman's (1969) factor analysis of the Adjective Check List (Gough & Heilbrun, 1965). Like the PC instrument, it asks how often each characteristic is true of the real self, ideal self, and ought self. RI is computed as the average absolute difference between the real-self and ideal-self ratings across all characteristics; RO is computed in parallel fashion (see the Method section). This instrument is presumably less meaningful to the individual than is the PC instrument, because its content is a uniform set of characteristics, but it has the advantage of faster administration. In the development of this instrument, a study using the RI measure supported its test–retest reliability ($r = .76$ over 3 to 5 weeks) and its test-criterion relationship to neuroticism (Watson & Watts, 2001).

AM

Like the PC instrument, the AM instrument was designed to correspond to the experience of the individual. This content-free instrument is a modification of Shlien's (1962) Abstract Apparatus, which was based on his clinical observation that the person has direct access to the experience of self-discrepancy. The participant provides a single rating of RI and a single rating of RO reflecting the participant's perception of the magnitude of these discrepancies. Specifically, the participant indicates in general how much alike are the real self and ideal self and how much alike are the real self and ought self by selecting a pair of intersecting squares (Study 1) or circles (Study 2) that represents the degree of similarity. The discrepancy score is computed by subtracting the proportion of the intersection from 1 (see the Method section). The use of a single item for each discrepancy permits very fast administration, though it may compromise the test–retest reliability and the validity of the scores.

We hypothesized that the PC instrument would generally have the strongest psychometric properties because, unlike the CC instrument, it has meaningful, idiographic content and because, unlike the single-item AM, it has multiple items.

Use of Difference Scores in Discrepancy-Between-Perceptions Instruments

Difference scores have been criticized because they bear the errors of the two variables that comprise them, negatively affecting reliability and validity (e.g., Bereiter, 1963; Cronbach & Furby, 1970; Linn & Slinde, 1977; Lord, 1963; Wylie, 1974). In her discussion of the measurement of self-concept, Wylie (1974) recommended use of perception-of-discrepancy instruments as one way of addressing the criticisms of difference scores. Regarding discrepancy-between-perceptions instruments, Wylie disagreed with the recommendation by other critics to keep the real-self and ideal-self component scores separate in a multivariate approach. She made a basic distinction between a difference score for self-discrepancy and other types of difference scores by arguing that, unlike other constructs for which difference scores have been used, self-discrepancy is an experiential entity, "a difference between two 'points,' both by definition in the phenomenal field of the *S*" (Wylie, 1974, p. 89). Nevertheless, Wylie questioned the reliability and validity of difference scores because they bear the errors of two variables, so she recommended using the real-self component score instead of a difference score. However, in a study that used an earlier version of the PC-RI measure used in the present study, Watson and Watts (2001) found that the test–retest reliability coefficients of the RI score and the real-self score were very similar and that RI scores predicted unique variance in the criterion neuroticism whereas real-self scores did not. Those findings support the use of a difference score rather than a real-self component score. In the present study, preliminary analyses similar to the analyses reported by Watson and Watts were conducted to confirm that the discrepancy-between-perceptions instruments yield difference-score self-discrepancy variables that outperform the variables that comprise them.

Evaluation of Psychometric Properties

In evaluating the psychometric properties of the three instruments, we examined internal consistency coefficients, test–retest reliability coefficients, convergent and discriminant evidence of validity, and predictive validity of test-criterion relationships (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), as well as invariance of the instruments' measurement properties across time and across samples. Following Kline's (2005) approach, we used confirmatory factor analysis (CFA) to examine convergent and discriminant validities. CFA is widely recommended for this purpose because it uses contemporary modeling techniques to implement the original logic of Campbell and Fiske's (1959) multitrait–multimethod matrix for validating convergent and discriminant validities (Brown, 2006; Kline, 2005). We also used CFA to evaluate measurement invariance, test–retest reliability coefficients, and test-criterion relationships.

To evaluate the test-criterion relationships of the scores, we based the criteria on consideration of both Rogers's (1959) theory and Higgins's (1987) theory. Although Rogers did not theorize explicitly about the RO construct, a high RO corresponds closely to Rogers' concept of the introjection of conditions of worth from significant others. Rogers's theory implies that a high RO leads to a high RI, which is a vulnerability to emotional distress. According to this theory, both anxiety and depression are criteria for both discrepancies, with RI being the more proximal predictor, which is supported by previous findings (Bryan, Watson, Babel, & Thrash, 2008). According to Higgins's theory that RO is uniquely related to anxiety and RI is uniquely related to depression, anxiety is a more relevant criterion for RO, and depression is more relevant for RI. There has been inconsistent support for Higgins's theory in eight studies that have tested both hypotheses, using clinical samples or clinical measures with nonclinical samples and following Strauman, Vookles, Berenstein, Chaiken, and Higgins's (1991) guidelines to control the alternate discrepancy and the alternate emotional distress. Three of these studies supported both hypotheses (Strauman, 1989, 1992; Strauman & Higgins, 1988, Study 2), whereas two found the opposite of both hypotheses (Tangney et al., 1998; Weilage & Hope, 1999), one found the opposite of the RO hypothesis and did not support the RI hypothesis (Bryan et al., 2008), one supported only the RI hypothesis (Bruch et al., 2000), and one supported neither hypothesis (Scott & O'Hara, 1993). Apart from inconsistent support for Higgins's theory, studies have provided evidence of the associations of both discrepancies to both anxiety and depression (e.g., Tangney et al., 1998; Weilage & Hope, 1999). In the present studies, we used clinical measures of both anxiety and depression to evaluate the test-criterion relationships of both RI and RO.

Overview of the Present Studies

We evaluated and compared the psychometric properties of the three self-discrepancy instruments, using a clinical sample and a nonclinical sample so that the findings would be applicable to clinical research and practice and to personality and social psychology research. In Study 1, the clinical sample was used to evaluate internal consistency coefficients, convergent and discriminant evidence of validity, and validity of test-criterion relationships at one time point with one measure of anxiety and one measure of depression as criteria. In Study 2A, the nonclinical sample was used to evaluate measurement invariance across a 4-week period, test–retest reliability coefficients, predictive validity of test-criterion relationships with three measures each of anxiety and depression as criteria, and, again, internal consistency coefficients and convergent and discriminant evidence of validity. Response bias was controlled in the analyses in Study 2A. With the data from both the clinical and the nonclinical samples, in Study 2B we tested measurement invariance across groups, and we used group membership as a second validity criterion in addition to scores on anxiety and depression measures in Studies 1 and 2A. Our general hypothesis across studies, as stated earlier, was that the PC instrument would have the strongest psychometric properties because it has multiple items and meaningful, idiographic content.

Study 1

Method

Participants. Ninety-nine female and 26 male undergraduate and graduate students seeking individual therapy at a university clinic volunteered to participate for remuneration. Ages ranged from 18 to 37 years, with a mean of 21.2 ($SD = 3.0$). Ethnicity was 83.2% White Americans. Among participants, 13.6% were taking medication for depression, and 8.0% were taking medication for anxiety.

Instruments.

Self-discrepancy instruments.

SCQ-PC. The idiographic SCQ-PC, an online computer program, measures real self, ideal self, and ought self with bipolar personal constructs. The participant lists six characteristics that describe the real self ("yourself as YOU see yourself in your own eyes"), six characteristics that describe the ideal self ("yourself as YOU would like to be in your own eyes"), and six characteristics that describe the ought self ("yourself as OTHERS think you ought or should be"). These 18 characteristics then are presented in a random order, and the participant enters the opposite of each characteristic, yielding an additional six characteristics for each self-component. Thus, the procedure elicits six characteristics and six opposite characteristics for each self-component, generating 12 characteristics for each of the three self-components. The 36 characteristics then are presented in a random order, and the participant rates them on a scale from 1 (*never or almost never true*) to 7 (*always or almost always true*) in response to each of the definitions of real self, ideal self, and ought self that were used to elicit the characteristics.

The SCQ-PC RI discrepancy (PC-RI) is scored by calculating the absolute difference between the real-self rating and the ideal-self rating on each of the 12 characteristics of the real self and on each of the 12 characteristics of the ideal self and then calculating the mean of the 24 absolute difference scores. The absolute difference indicates the distance between the real-self rating and the ideal-self rating for a characteristic whether the ideal-self rating is positive or negative. The SCQ-PC RO discrepancy (PC-RO) is scored in a similar way with the 12 characteristics of the real self and the 12 characteristics of the ought self.

SCQ-CC. This online questionnaire measures real self, ideal self, and ought self, using 28 personality characteristics identified by Parker and Veldman (1969) as the four highest loading items on each of seven factors in a factor analysis of the Adjective Check List (Gough & Heilbrun, 1965). The participant rates the 28 characteristics in response to the same instructions and on same the scale used for the SCQ-PC.

The SCQ-CC RI discrepancy (CC-RI) is scored by calculating the absolute difference between the real-self and ideal-self ratings on each of the 28 characteristics and then calculating the mean of the 28 absolute difference scores. The SCQ-CC RO discrepancy (CC-RO) is scored in a similar way with the real-self and ought-self ratings.

AM. The AM measures RI and RO discrepancies. It is an online, modified version of Shlien's (1962) content-free Abstract Apparatus that assesses self-discrepancy by using nine pairs of squares, the areas of which intersect 0%, 12.5%, 25%, 37.5%, 50%, 62.5%, 75%, 87.5%, and 100%. To measure the RI discrepancy (AM-RI), real self and ideal self are defined in the same way as in the SCQ-PC and SCQ-CC, and the participant is asked to think of one square as representing the real self and the other square as representing the ideal self. The participant then selects the pair of squares with the intersecting area that "shows how much your real self and ideal self are alike in general." The discrepancy score is calculated by subtracting the proportion of the intersecting area from 1. The RO discrepancy (AM-RO) is measured with a similar procedure.

Anxiety and depression instruments. Anxiety was assessed with the State–Trait Anxiety Inventory–Trait scale–Anxiety factor (STAI-T-A). The Trait scale of the STAI (Spielberger, 1983) has been factored to yield a seven-item Anxiety factor that shows stronger discriminant validity in relation to measures of depression than does the full Trait scale (Bieling, Antony, & Swinson, 1998; Vautier, Callahan, Moncany, & Sztulman, 2004). Depression was assessed with the Beck Depression Inventory–II (BDI-II; Beck, Steer, & Brown, 1996).

Procedure. Participants completed the discrepancy instruments first in partially counterbalanced orders and the anxiety and depression instruments second in counterbalanced orders. The orders of real self, ideal self, and ought self were counterbalanced.

Results

Preliminary analyses. All observed variables were normally distributed. Descriptive statistics and internal consistencies are shown in Table 1. Intercorrelations of observed variables are shown in Table 2.

We conducted preliminary analyses of the discrepancy-between-perceptions instruments to address the criticism that the two error terms in a discrepancy score make its validity weaker than that of the real-self score alone (e.g., Wylie, 1974). For PC-RI, PC-RO, CC-RI, and CC-RO, we tested the hypothesis that the discrepancy score predicts unique variance in the criteria STAI-T-A and BDI-II but that the real-self score does not. We tested a similar hypothesis for the discrepancy score and the ideal-self score as predictors and for the discrepancy score and the ought-self score as predictors. Real-self and ideal-self scores were computed, reverse scoring items for which the rating of the ideal self was lower than that of the real self or for which the ratings of both self-components were the same and below the midpoint of the scale. The same procedure was used to score the real self and ought self.

To test the incremental validity of discrepancy and real-self scores, PC-RI and the corresponding real-self variable were entered as simultaneous predictors of BDI-II and STAI-T-A. Parallel sets of analyses were conducted for PC-RO, CC-RI, and CC-RO. The hypothesis that the discrepancy score would predict unique variance in the criterion and that the real-self score would not was supported in seven of the eight analyses (ps $< .05$); the exception was that neither CC-RO nor CC real self predicted unique variance in BDI-II. Regarding the incremental validity of discrepancy and ideal-self scores, the hypothesis that the discrepancy score would predict unique variance in the criterion and that the ideal-self score would not was supported in all four analyses (ps $< .001$). Regarding the incremental validity of discrepancy and ought-self scores, the hypothesis that the discrepancy score would predict unique variance in the criterion and that the ought-self score would not

Table 1
Descriptive Statistics and Internal Consistencies for Study 1 (N = 125) and Study 2A (N = 278) Observed Variables

Measure	*M*	*SD*	Range	α	AIC
Study 1					
PC-RI	2.08	1.04	0.17–4.70	.92	.34
PC-RO	2.02	0.93	0.17–4.09	.90	.29
CC-RI	1.72	0.63	0.50–3.54	.82	.19
CC-RO	1.78	0.59	0.36–3.61	.84	.16
AM-RI[a]	0.55	0.20	0.125–1.00	—	—
AM-RO[a]	0.58	0.17	0.25–0.88	—	—
BDI-II	17.90	9.94	0–42	.90	.30
STAI-T-A	17.52	4.75	7–28	.83	.42
Study 2A					
PC-RI T1	1.47	0.82	0.04–4.70	.91	.31
PC-RI T2	1.41	0.82	0.00–4.78	.92	.33
PC-RO T1	1.33	0.77	0.26–3.70	.90	.28
PC-RO T2	1.38	0.81	0.00–4.65	.91	.31
CC-RI T1	1.40	0.58	0.32–3.46	.88	.20
CC-RI T2	1.37	0.63	0.18–3.75	.90	.24
CC-RO T1	1.43	0.62	0.21–3.50	.89	.22
CC-RO T2	1.42	0.67	0.14–3.50	.90	.25
AM-RI T1[a]	0.45	0.19	0.00–1.00	—	–
AM-RI T2[a]	0.45	0.19	0.00–1.00	—	–
AM-RO T1[a]	0.45	0.19	0.00–0.84	—	–
AM-RO T2[a]	0.47	0.19	0.00–1.00	—	–
BDI-II	7.72	7.83	0–41	.88	.34
CES-D	31.69	9.57	20–68	.91	.33
DASS-D	4.69	6.57	0–39	.91	.51
BAI	28.22	7.70	21–79	.92	.32
STAI-T-A	12.99	4.08	7–26	.84	.40
DASS-A	3.76	4.91	0–35	.85	.31
IM T1	7.10	3.60	0–18	.75	.13
IM T2	6.99	3.92	0–19	.79	.16

Note. AIC = average interitem correlation; PC = Personal Constructs; CC = Conventional Constructs; AM = Abstract Measure; RI = real–ideal discrepancy; RO = real–ought discrepancy; BDI-II = Beck Depression Inventory–II; CES-D = Center for Epidemiologic Studies–Depression Scale; DASS-D = Depression Anxiety Stress Scales–Depression scale; BAI = Beck Anxiety Inventory; STAI-T-A = State–Trait Anxiety Inventory–Trait scale–Anxiety factor; DASS-A = Depression Anxiety Stress Scales–Anxiety scale; IM = impression management; T1 = Time 1; T2 = Time 2.
[a] The AMs have only one item, so internal consistency and interitem correlation are not applicable.

was supported in three of the four analyses (*p*s $< .001$); the exception was that neither PC-RO nor PC ought self predicted unique variance in STAI-T-A. These results show that the discrepancy score generally has incremental test-criterion validity but that the real-self, ideal-self, and ought-self scores do not. These analyses indicate that it would be undesirable to replace the difference-score-based discrepancy variables with real-self, ideal-self, or ought-self component scores.

Overview of modeling strategy. In this study and the following ones, primary analyses were conducted with Amos 7.0

Table 2
Study 1: Intercorrelations of Self-Discrepancy, Depression, and Anxiety Observed Variables (N = 125)

Measure	1	2	3	4	5	6	7	8
1. PC-RI	—							
2. PC-RO	.74***	—						
3. CC-RI	.73***	.60***	—					
4. CC-RO	.46***	.65***	.69***	—				
5. AM-RI	.68***	.53***	.65***	.49***	—			
6. AM-RO	.29**	.38***	.35***	.39***	.46***	—		
7. BDI-II	.53***	.36***	.57***	.32***	.59***	.21*	—	
8. STAI-T-A	.36***	.21*	.52***	.35***	.46***	.08	.59***	—

Note. PC = Personal Constructs; CC = Conventional Constructs; AM = Abstract Measure; RI = real–ideal discrepancy; RO = real–ought discrepancy; BDI-II = Beck Depression Inventory–II; STAI-T-A = State–Trait Anxiety Inventory–Trait scale–Anxiety factor.
* $p < .05$. ** $p < .01$. *** $p < .001$.

(Arbuckle, 2006) with maximum likelihood estimation. Covariance matrices served as input. Unless otherwise stated, latent variables in each model were specified to be correlated. Model fit was evaluated with three widely used fit indexes: the Tucker–Lewis index (TLI), the comparative fit index (CFI), and the root-mean-square error of approximation (RMSEA). TLI and CFI values close to 1 and RMSEA values close to 0 indicate good fit. We followed Hu and Bentler's (1999) recommendations in using the following cutoff criteria: TLI, .95; CFI, .95, and RMSEA, .06. The relative fits of nested models, in which one model is a constrained version of another model, were compared directly with a chi-square difference test. A significant increase in chi-square as parameters are constrained indicates a significant decline in fit (Bollen, 1989).

Convergent and discriminant evidence of validity. Fit indexes for models tested in this study are reported in Table 3. We examined convergent and discriminant evidence of validity with a correlated uniqueness model, a contemporary and widely recommended CFA-based approach to analysis of multitrait–multimethod data (Brown, 2006; Kenny & Kashy, 1992; Marsh, 1989). Constructs (traits) are modeled as latent factors, and shared method variance is removed by modeling correlations between the error terms of observed variables assessed by the same instrument (method). The model had two latent factors: RI Discrepancy and RO Discrepancy. PC-RI, CC-RI, and AM-RI were specified as indicators of the RI factor; PC-RO, CC-RO, and AM-RO were specified as indicators of the RO factor. To account for method variance related to particular instruments, correlations were modeled among the residual error terms of the PC observed variables and likewise among the residual error terms of the CC and AM observed variables. The model demonstrated good fit, as shown in Table 3 (Model 1). All loadings were significant ($p < .001$; see the Convergent Evidence of Validity section for loadings). The RI and RO latent variables were strongly correlated, $r = .77$, $p < .001$, raising the question of whether RI and RO are distinct latent constructs.

Discriminant evidence of validity. Following Kline (2005), we examined discriminant evidence of validity of the RI and RO latent variables by comparing the fit of the two-factor model with that of a more constrained one-factor model (Table 3, Model 2). The one-factor model fit significantly worse than the two-factor model, $\chi^2_{\text{diff}}(1) = 51.91$, $p < .001$, and had unacceptable fit (see Hu & Bentler's, 1999, criteria described earlier). These findings indicate that the RI and RO latent variables, although correlated, are factorially distinct, providing discriminant evidence of validity of the latent variables.

Having found that RI and RO are distinct latent constructs, we next tested whether particular instruments show discriminant evidence of validity in assessing these constructs. The good fit of the two-factor CFA suggests, but does not guarantee, that each measure of RI and RO assessed only the discrepancy variable that it was designed to measure. To address this issue directly, we conducted six separate tests allowing PC-RI, PC-RO, CC-RI, CC-RO, AM-RI, and AM-RO to cross-load on the alternate discrepancy latent variable. These tests yielded no proper solutions with significant cross-loadings. These results establish discriminant evidence of validity for all the individual measures.

Convergent evidence of validity. Following Kline (2005), we addressed convergent evidence of validity of the PC, CC, and AM measures by examining whether each observed variable loaded robustly on the discrepancy it was designed to measure. Standardized loadings in Model 1 (Table 3) were strong for PC-RI, CC-RI, AM-RI, PC-RO, and CC-RO (.76–.87) but were weak for AM-RO (.46). Using the z test recommended by Meng, Rosenthal, and Rubin (1992), we tested the significances of the differences among the loadings. On the RI latent variable, PC-RI loaded more strongly than did AM-RI ($z = 2.00$, $p < .05$). On the RO latent variable, PC-RO loaded more strongly than did AM-RO ($z = 5.27$, $p < .001$), and CC-RO loaded more strongly than did AM-RO ($z = 3.49$, $p < .001$).[2] In summary, convergent evidence of validity was weak for AM-RO but strong for all other discrepancy measures.

Test-criterion evidence of validity. To evaluate test-criterion relationships, we examined the correlations of the discrepancy variables with the anxiety and depression (distress) variables. Correlations of the observed variables all were significant except AM-RO with anxiety (see Table 2). However, correlations between observed variables are ambiguous, because an observed discrepancy variable could be associated with the criterion because (a) it loads on the underlying latent variable, which in turn is related to the criterion, or (b) it contains error variance that is related to the criterion (for details, see Kline's, 2005, tracing rules). To address the ambiguity of the observed correlations, we used CFA to model both sources of covariance.

In an initial four-factor CFA, anxiety and depression latent variables were added to the two-factor CFA model described earlier. Three random parcels of STAI-T-A items were modeled as indicators of the anxiety latent variable, and three random parcels of BDI items were modeled as indicators of the depression latent variable. All parcels were normally distributed. The model had good fit (Table 3, Model 3). Both the RI and RO discrepancy latent variables were significantly positively correlated with both the depression and anxiety latent variables (RI and depression, $r = .69$; RI and anxiety, $r = .66$; RO and depression, $r = .47$; RO and anxiety, $r = .23$; all $p\text{s} < .001$), supporting the test-criterion relationships of the discrepancy constructs. Accordingly, the test-criterion evidence of validity for the observed variables was partially or entirely attributable to these relationships among underlying latent constructs.

Next, we tested whether the test-criterion evidence of validity for the observed variables was due in part to error in the observed discrepancy variables. In 12 separate analyses, the error term of one of the six observed discrepancy variables was specified to correlate with one of the two distress latent variables. Two of these correlations were significant and were therefore retained (Table 3, Model 4; see Figure 1). The CC-RO error term was positively correlated with the anxiety latent variable, $r = .28$, $p < .01$, indicating that the correlation of the observed CC-RO with the

[2] The finding that the PC and CC measures loaded more strongly than the AM could be an artifact of the similarity of their method of measurement. To examine this possibility in both this study and Study 2A, method variance unique to PC and CC was modeled by specifying correlations between the error terms of the PC-RI and CC-RI observed variable and between the error terms of the PC-RO and CC-RO observed variable. These correlations were not significant, indicating that the higher loadings of PC and CC measures are not artifacts of an unmodeled source of shared method variance.

Table 3
Model Goodness-of-Fit Indicators for Study 1 (N = 125)

Model	*df*	χ^2	TLI	CFI	RMSEA
Model 1: Two factors	5	5.43	1.00	1.00	.03
Model 2: One factor	6	57.33***	.71	.88	.26
Model 3: Criterion validity without correlated errors	45	58.39	.98	.99	.05
Model 4: Criterion validity with correlated errors	43	44.26	1.00	1.00	.02

Note. TLI = Tucker–Lewis index; CFI = comparative fit index; RMSEA = root-mean-square error of approximation.
*** $p < .001$.

anxiety variable is inflated by residual error variance. The PC-RI error term was negatively correlated with the anxiety latent variable, $r = -.31$, $p < .001$, indicating that the correlation of the observed PC-RI with the anxiety variable is reduced by residual error variance.

PC-RO, CC-RI, and AM-RI had strong test-criterion relationships in that their observed correlations with the criteria were due to their strong convergence on the discrepancy latent variables, which relate to the criteria, and were not due to correlations of their error terms with the criteria. PC-RI also had strong test-criterion relationships in that it loaded strongly on the latent variable. Its observed correlation with anxiety was an underestimate because of the negative correlation of its error with anxiety. However, CC-RO had a weak test-criterion relationship in that its observed correlation with anxiety was inflated because of the positive correlation of its error with anxiety. AM-RO had unacceptable test-criterion relationships in that it failed to load strongly on the RO latent variable, resulting in a nonsignificant observed correlation with anxiety (see Table 2).

Discussion

For the clinical sample in Study 1, preliminary analyses of the discrepancy-between-perceptions instruments supported the use of the discrepancy scores, showing that they generally predicted unique variance in the anxiety and depression scores but that the real-self, ideal-self, and ought-self scores did not. Internal consistency alpha coefficients were strong for PC-RI (.92) and PC-RO (.90) but were moderate for CC-RI (.82) and CC-RO (.84), and average interitem correlations all fell within Clark and Watson's (1995) ideal range of .15 to .50. Although strongly correlated, the RI and RO latent variables were found to be factorially distinct. Convergent evidence of validity was strong for all measures except AM-RO, for which it was weak. Convergence was stronger for both of the PC measures than for their respective AM, but for the CC measures, convergence was stronger than that of the AM only for RO. Discriminant evidence of validity was found for all the individual RI and RO measures. Test-criterion evidence of validity at one time point, using one clinical measure of anxiety and one

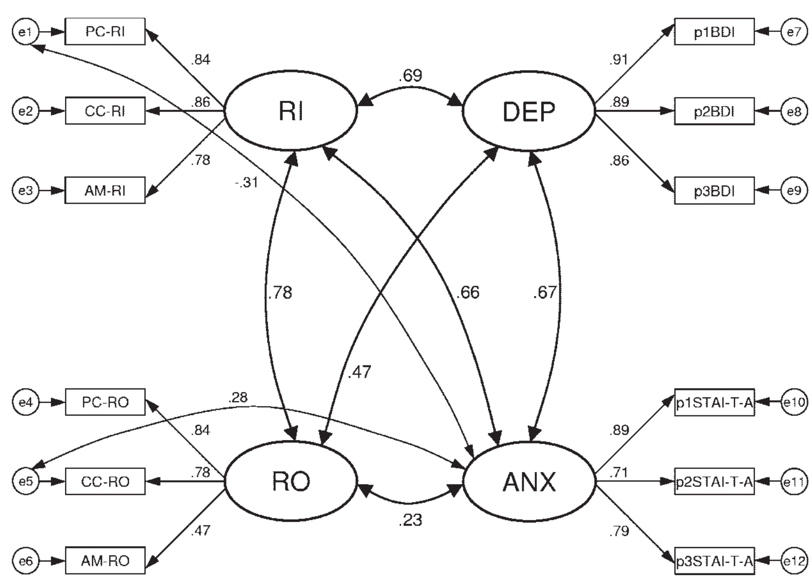

Figure 1. Study 1, Table 3, Model 4: Four-factor confirmatory factor analysis. Error correlations were excluded from the figure for presentation clarity. All coefficients were significant at $p < .001$. e1–e12 = Error Terms 1–12; RI = real–ideal discrepancy; RO = real–ought discrepancy; DEP = depression; ANX = anxiety; PC = Personal Constructs; CC = Conventional Constructs; AM = Abstract Measures; BDI = Beck Depression Inventory–II; STAI-T-A = State–Trait Anxiety Inventory–Trait scale–Anxiety factor; p1–p3 = Parcels 1–3.

clinical measure of depression as criteria, was found for the RI and RO latent variables and for PC-RI and PC-RO, CC-RI, and AM-RI; however, it was weak for CC-RO, and AM-RO was not correlated with anxiety. In summary, comparisons of the psychometric properties of the three instruments showed that the properties of the idiographic PC measures were consistently as strong as or stronger than those of the CC measures and the AM.

Study 2A

In Study 2A, we repeated with a nonclinical student sample the tests in Study 1 and added analyses of measurement invariance and test–retest reliability across a 4-week period. Predictive validity of test-criterion relationships over the same time period was evaluated with three measures each of anxiety and depression. Effects of impression management response bias were controlled as recommended by Podsakoff, MacKenzie, Lee, and Podsakoff (2003).

Method

Participants. One hundred thirty-seven female and 141 male students participated for course credit in an introductory psychology course. Ages ranged from 18 to 40 years, with a mean of 19.5 ($SD = 2.1$). Ethnicity was 78.1% White Americans. Among participants, 3.2% were in therapy, 4.0% were taking medication for depression, and 3.2% were taking medication for anxiety.

Instruments.

Self-discrepancy instruments. The same PC and CC instruments as in Study 1 were used. On the AM instrument, the scaling was changed from nine pairs of intersecting squares to seven pairs of circles, the areas of which intersect 0%, 16.66%, 33.33%, 50%, 66.66%, 83.33%, and 100%; in other respects the AM instrument was the same as in Study 1.

Anxiety and depression instruments. Anxiety was assessed with the STAI-T-A, as in Study 1, and also with the Beck Anxiety Inventory (BAI; Beck & Steer, 1990) and the Anxiety scale of the Depression Anxiety Stress Scales (DASS-A; Lovibond & Lovibond, 1993).

Depression was assessed with the BDI-II, as in Study 1, and also with the Center for Epidemiologic Studies–Depression Scale (CES-D; Radloff, 1977) and the Depression scale of the DASS (DASS-D; Lovibond & Lovibond, 1993).

Response bias instrument. Response bias was measured with the Impression Management scale of the Balanced Inventory of Desirable Responding (Paulhus, 1984).

Procedure. Participants completed the discrepancy instruments at Time 1. Four weeks later at Time 2, participants completed instruments in this order: discrepancy, impression management, and anxiety and depression. The discrepancy instruments were partially counterbalanced, as were the anxiety and depression instruments. Real, ideal, and ought selves were counterbalanced. Of the Time 1 sample, 93% completed the study; data from noncompleters were not analyzed.

Results

Preliminary analyses. Comparisons of the attrition group ($N = 21$) with the final sample ($N = 278$), using t tests on all observed variables, showed no significant differences on Time 1 variables except that the final sample was higher on impression management, $t(297) = 3.72$, $p < .001$.

Descriptive statistics and internal consistencies for measures at Time 1 and Time 2 are presented in Table 1. Intercorrelations of the observed variables are presented in Table 4.

None of the discrepancy variables was significantly skewed and only two of 12 observed discrepancy variables had mild kurtosis at Time 2. Discrepancy variables were left in their raw metric so that tests of invariance across time would not be biased and because maximum likelihood estimation is robust against mild violations of normality (McDonald & Ho, 2002). However, three of the six anxiety and depression variables had significant skewness, and four of six had significant kurtosis. We transformed these variables, using a square root function, yielding normal distributions.

To address the criticism that the two error terms in a discrepancy score make it less reliable than its two components (e.g., Wylie, 1974), for the PC and CC instruments we compared the test–retest reliabilities of the discrepancy scores to those of the real-self, ideal-self, and ought-self scores. Real-self, ideal-self, and ought-self scores were computed as described in Study 1. For the RI measures, reliabilities were .72 for PC-RI, .73 for PC real self, and .52 for PC ideal self; reliabilities were .77 for CC-RI, .78 for CC real self, and .59 for CC ideal self. For the RO measures, reliabilities were .74 for PC-RO, .63 for PC real self, and .54 for PC ought self; reliabilities were .72 for CC-RO, .76 for CC real self, and .66 for CC ought self. These findings support the test–retest reliabilities of the difference scores, which were generally at least as strong as those of their components.

Convergent and discriminant evidence of validity. Fit indexes for models tested in this study are reported in Table 5. We examined convergent and discriminant evidence of validity, using the same analytic strategy as in Study 1. The CFA model had six factors: RI discrepancy, RO discrepancy, and impression management at both Time 1 and Time 2. At both Time 1 and Time 2, PC-RI, CC-RI, and AM-RI were specified as indicators of RI; PC-RO, CC-RO, and AM-RO were specified as indicators of RO; three parcels of randomly selected items were specified as indicators of impression management. All discrepancy latent variables were specified to covary. We removed the effect of impression management from each observed variable, using a modeling strategy recommended by Podsakoff et al. (2003; see Figure 2). Autocorrelations were modeled between corresponding error terms at Time 1 and Time 2 (Pitts, West, & Tein, 1996). At Time 1 and Time 2, indicators of RI and RO assessed by the same method were specified to have correlated error terms. The model demonstrated good fit (Table 5, Model 1). All discrepancy loadings were significant ($ps < .001$). As expected, loadings of the observed discrepancy variables on impression management were negative ($-.24$ to $-.03$). RI and RO latent variables were strongly correlated at Time 1, $r = .76$, $p < .001$, and Time 2, $r = .83$, $p < .001$.

Discriminant evidence of validity. To examine discriminant evidence of validity of the RI and RO latent variables, we compared the fit of the six-factor model (Table 5, Model 1) to the fits of two five-factor models in which the Time 1 or Time 2 indicators of RI and RO were specified to load on a single factor (Table 5, Models 2 and 3). Both five-factor models fit significantly worse than the six-factor model, $\chi^2_{\text{diff}}(1) \geq 71.98$, $p < .001$, and both had unacceptable fit. These findings indicate that the RI and RO latent variables represent distinct latent constructs.

Table 4
Study 2A: Intercorrelations of Self-Discrepancy, Depression, and Anxiety Observed Variables (N = 278)

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. PC-RI Time 1	—																			
2. PC-RO Time 1	.67	—																		
3. CC-RI Time 1	.70	.56	—																	
4. CC-RO Time 1	.51	.73	.62	—																
5. AM-RI Time 1	.56	.37	.48	.27	—															
6. AM-RO Time 1	.36	.44	.35	.36	.40	—														
7. PC-RI Time 2	**.72**	.55	.62	.45	.47	.29	—													
8. PC-RO Time 2	.58	**.74**	.50	.64	.35	.39	.74	—												
9. CC-RI Time 2	.64	.52	**.77**	.56	.47	.34	.67	.58	—											
10. CC-RO Time 2	.52	.65	.60	**.72**	.32	.38	.55	.72	.75	—										
11. AM-RI Time 2	.47	.29	.40	.23	**.64**	.39	.53	.42	.50	.37	—									
12. AM-RO Time 2	.44	.43	.32	.34	.44	**.59**	.44	.48	.44	.42	.58	—								
13. BDI-II	**.40**	**.27**	**.38**	**.23**	**.40**	**.23**	.48	.39	.42	.37	.47	.33	—							
14. CES-D	**.45**	**.31**	**.41**	**.31**	**.42**	**.24**	.58	.47	.50	.43	.49	.36	.84	—						
15. DASS-D	**.39**	**.26**	**.33**	**.23**	**.35**	**.19**	.57	.48	.44	.40	.43	.32	.82	.87	—					
16. BAI	**.24**	**.16**	**.22**	**.20**	**.18**	**.09**	.31	.25	.24	.25	.22	.15	.58	.61	.53	—				
17. STAI-T-A	**.37**	**.27**	**.38**	**.32**	**.33**	**.16**	.43	.37	.41	.40	.31	.23	.63	.68	.61	.55	—			
18. DASS-A	**.29**	**.19**	**.26**	**.24**	**.23**	**.12**	.37	.31	.31	.25	.24	.19	.63	.68	.63	.83	.60	—		
19. IM Time 1	−.13	−.07	−.19	−.04	−.19	−.22	−.20	−.16	−.14	−.09	−.16	−.16	−.24	−.29	−.24	−.20	−.22	−.18	—	
20. IM Time 2	−.09	−.02	−.14	−.02	−.18	−.16	−.22	−.15	−.12	−.08	−.11	−.08	−.25	−.24	−.21	−.18	−.25	−.20	.81	—

Note. Correlations in bold are self-discrepancy test–retest and criterion validity correlations. PC = Personal Constructs; CC = Conventional Constructs; AM = Abstract Measure; RI = real–ideal discrepancy; RO = real–ought discrepancy; BDI-II = Beck Depression Inventory–II; CES-D = Center for Epidemiologic Studies–Depression Scale; DASS-D = Depression Anxiety Stress Scales–Depression scale; BAI = Beck Anxiety Inventory; STAI-T-A = State–Trait Anxiety Inventory–Trait scale–Anxiety factor; DASS-A = Depression Anxiety Stress Scales–Anxiety scale; IM = impression management. If r = .13–.15, $p < .05$. If r = .16–.20, $p < .01$. If $r \geq .21$, $p < .001$. If $r \leq .12$, p = *ns.*

Next, we examined discriminant evidence of validity of the individual measures PC-RI, PC-RO, CC-RI, CC-RO, AM-RI, and AM-RO, using the same analytic strategy as in Study 1. Allowing Time 1 PC-RI to load on the Time 1 RO latent variable and Time 2 PC-RI to load on the Time 2 RO latent variable resulted in no significant cross-loadings. Parallel tests revealed that allowing PC-RO, CC-RI, CC-RO, and AM-RI to cross-load on the alternate discrepancy failed to result in significant cross-loadings. In contrast, a parallel test for AM-RO revealed a significant cross-loading at Time 2, loading = .31, $p < .01$. The resulting model (Table 5, Model 4) is shown in Figure 2. These results indicate discriminant evidence of validity of all individual measures except AM-RO.

Convergent evidence of validity. As in Study 1, we examined convergent evidence of validity of the PC, CC, and AM individual measures of RI and RO by testing whether each observed variable loaded robustly on the discrepancy it was designed to measure. The standardized loadings of each of the three discrepancy measures in Model 4 (Table 5) are shown in Figure 2. Loadings were strong for PC-RI, PC-RO, CC-RI, and CC-RO (.80–.89), moderate for AM-RI (.60), and weak for AM-RO (.32–.47). We tested the significances of the differences among the loadings as in Study 1. For all four of the discrepancy latent variables, all three loadings differed from one another ($z \geq 2.22$, $p < .05$) except that the loadings of PC-RI and CC-RI were equal at Time 2 and the loadings of PC-RO and CC-RO were equal at Time 2.[3] On each latent variable at Time 1, the PC measure loaded more strongly than the CC measure, which loaded more strongly than the AM measure. On each latent variable at Time 2, the PC and CC measures loaded more strongly than the AM measure. In summary, convergent evidence of validity was strong for PC-RI, CC-RI, PC-RO, and CC-RO, moderate for AM-RI, but weak for AM-RO.

Temporal stabilities. We used means-and-covariance-structures analysis, an extension of CFA, to examine measurement invariance, individual differences in latent constructs (i.e., test–retest correlations), and measurement error. *Measurement invariance* refers to stability of the measurement model (e.g., loadings and intercepts), which relates observed variables to latent variables (Pitts et al., 1996). Stability of loadings and intercepts provides evidence that the observed variables assess the same construct and have a comparable metric at both time points. Establishing measurement invariance ensures a meaningful analysis of other aspects of stability (e.g., test–retest correlations). Because statistically significant levels of measurement noninvariance are sometimes negligible in a practical sense, Little (1997) recommended that the suitability of measurement invariance constraints be judged on the basis of a modeling rationale (specifically, $TLI_{diff} < .05$), rather than on the basis of a statistical rationale (i.e., a nonsignificant chi-square difference test). However, when testing substantive hypotheses about latent variable means or the relations of latent

[3] Meng et al.'s (1992) technique was used to test whether correlation coefficients are equal to one another. However, for AM-RO the loading on the RO latent variable is a standardized regression coefficient rather than a correlation coefficient because it cross-loads on the RI latent variable. To address this issue, we conducted the Meng test in models both with and without the cross-loaders modeled. In both cases the loadings of AM-RO were significantly weaker than PC-RO and CC-RO.

Table 5
Model Goodness-of-Fit Indicators for Study 2A (N = 278)

Model	*df*	χ^2	TLI	CFI	RMSEA
Model 1: Six factors	101	207.10***	.96	.97	.06
Model 2: Five factors, single discrepancy Time 1	102	289.78***	.92	.95	.08
Model 3: Five factors, single discrepancy Time 2	102	279.09***	.92	.95	.08
Model 4: AM-ROb as cross-loader	100	199.26***	.96	.97	.06
Model 5: Baseline model for invariance testing	99	198.10***	.96	.97	.06
Model 6: Invariance of loadings	104	205.98***	.96	.97	.06
Model 7: Invariance of intercepts	108	211.94***	.96	.97	.06
Model 8: Criterion validity without correlated errors	110	174.53***	.97	.98	.05
Model 9: Criterion validity with correlated errors	108	164.79***	.98	.98	.04

Note. TLI = Tucker–Lewis Index; CFI = comparative fit index; RMSEA = root-mean-square error of approximation; AM = Abstract Measure; RO = real–ought discrepancy; b = Time 2.
*** $p < .001$.

variables to one another, Little recommended using a statistical rationale. We followed Little's recommendations in this study and Study 2B, in which we tested measurement invariance across time and across groups, respectively.

Stability of factor loadings and intercepts (measurement invariance). In a test of the invariance of factor loadings across time, the baseline model was similar to Model 4 (Table 5) but included a means-and-intercepts structure. We identified the model by setting one loading on each latent variable equal to 1 and setting the corresponding intercepts equal to 0. This model had good fit (Table 5, Model 5). Constraining loadings and intercepts to be equal at Time 1 and Time 2 resulted in no loss of fit ($TLI_{diff} = .00$; Table 5, Models 6 and 7). The invariance of the loadings and intercepts provides evidence that the latent variables at Time 1 and Time 2 represent comparable constructs.

Test–retest correlations of the latent variables. On the basis of Model 7 (Table 5), the test–retest correlation was .92 for the RI latent variable and .91 for the RO latent variable (ps $< .001$). The correlations between the latent variables, from which error variance had been removed, were much higher than the correlations between the observed variables (see Table 4).

Stability of measurement error. On the basis of Model 7 (Table 5), the covariance between the Time 1 and Time 2 error terms was positive and significant for all measures except PC-RI and PC-RO. This finding suggests that the observed test–retest correlations for the CC and AM measures of discrepancy are inflated by stability of their error terms.

Test-criterion evidence of validity. To evaluate test-criterion relationships, we examined discrepancy variables at Time 1 as predictors of anxiety and depression variables at Time 2.

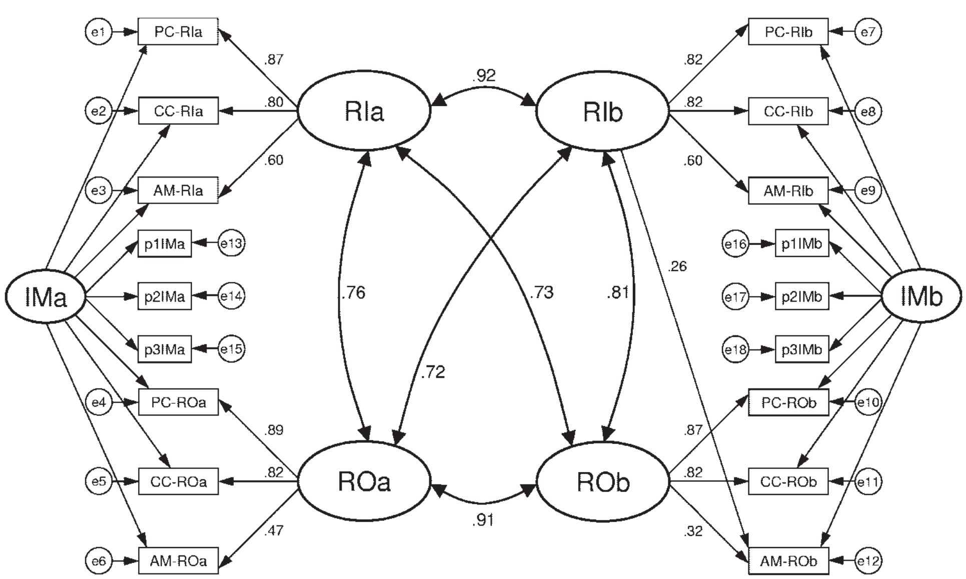

Figure 2. Study 2A, Table 5, Model 4: Six-factor confirmatory factor analysis with AM-ROb as a cross-loader. Error correlations were excluded from the figure for presentation clarity. The test–retest reliability of the impression management (IM) latent variable was also excluded, $r = .91$, $p < .001$. All coefficients were significant at $p < .001$. e1–e18 = Error Terms 1–18; RI = real–ideal discrepancy; RO = real–ought discrepancy; a = Time 1; b = Time 2; PC = Personal Constructs; CC = Conventional Constructs; AM = Abstract Measures.

Correlations of the observed variables all were significant except AM-RO with BAI and DASS-A (see Table 4). To address the ambiguity of correlations between observed variables (see Study 1), we tested a model that was similar to the initial six-factor CFA (Table 5, Model 1) except that Time 2 RI and RO latent variables were replaced with Time 2 depression and anxiety variables. Specifically, BDI, CES-D, and DASS-D were modeled as indicators of the depression latent variable, and BAI, STAI-T-A, and DASS-A were modeled as indicators of the anxiety latent variable. This model had good fit (Table 5, Model 8). Both the RI and RO discrepancy latent variables were significantly correlated with both the depression and anxiety latent variables (RI and depression, $r = .50$; RI and anxiety, $r = .36$; RO and depression, $r = .37$; RO and anxiety, $r = .24$; all ps $< .001$), indicating the criterion validity of the discrepancy constructs. Accordingly, observed variable correlations between discrepancy and criterion variables are partially or entirely attributable to these relationships among underlying latent constructs.

We next examined whether the test-criterion evidence of validity for the observed variables was due in part to error in the observed variables. In 12 separate analyses, the error term of one of the six discrepancy measures was specified to correlate with one of the two distress latent variables. Two of these correlations were significant and therefore were retained (Table 5, Model 9; see Figure 3). The AM-RI error term was correlated with the depression latent variable, $r = .09$, $p < .05$, and the CC-RO error term was correlated with the anxiety latent variable, $r = .11$, $p < .05$. Therefore, the observed correlations of AM-RI with the depression variables and the observed correlations of the CC-RO with the anxiety variables are inflated because they contain residual error variance that is related to the criterion variable. PC-RI, PC-RO, and CC-RI have acceptable test-criterion relationships in that their observed correlations with the criteria are due to their convergence on the discrepancy latent variables, which relate to the criteria, and are not due to correlations of their error terms with the criteria.

Discussion

For the nonclinical sample in Study 2A, preliminary analyses supported the use of discrepancy scores, showing that their test–retest reliability coefficients were generally at least as strong as those of the real-self, ideal-self, and ought-self scores. Internal consistency alpha coefficients were strong at both time points for PC-RI and PC-RO (.90 to .92) and for CC-RI and CC-RO (.88 to .90), and average interitem correlations all fell within Clark and Watson's (1995) ideal range of .15 to .50. The instruments demonstrated stable

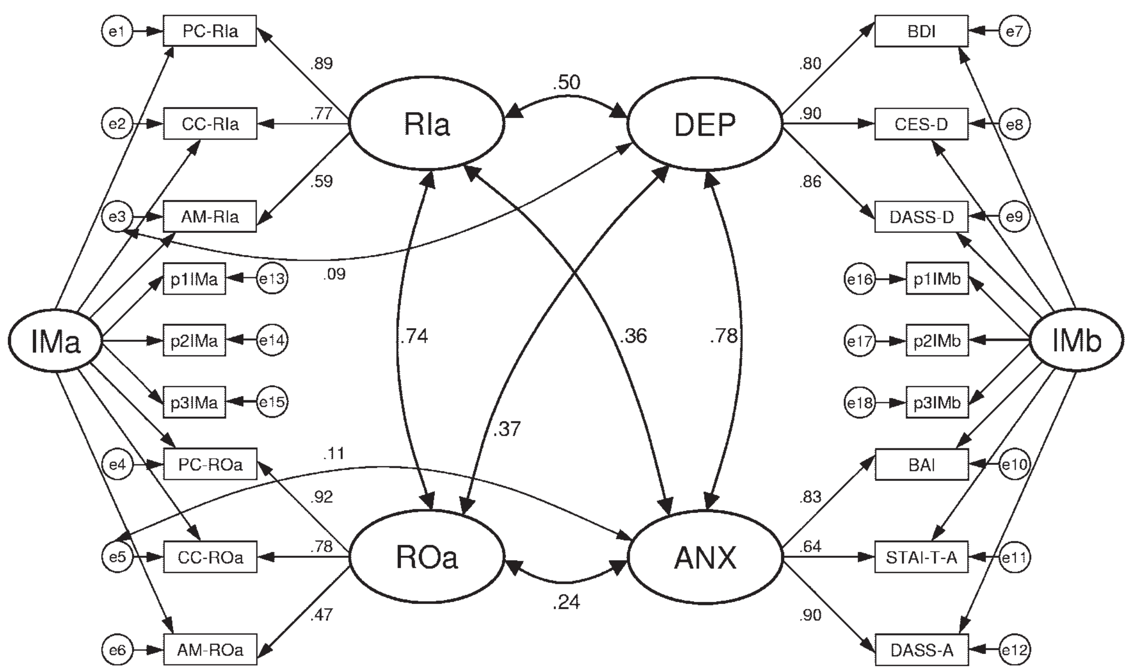

Figure 3. Study 2A, Table 5, Model 9: Six-factor confirmatory factor analysis criterion validity with self-discrepancy error terms correlated with distress latent variables. Error correlations were excluded from the figure for presentation clarity. The test–retest reliability of the impression management (IM) latent variable was also excluded for presentation clarity, $r = .91$, $p < .001$. All coefficients were significant at $p < .001$ with two exceptions: The correlation between e3 and the depression latent variable was significant at $p < .05$, and the correlation between e5 and the anxiety latent variable was significant at $p < .05$. e1–e18 = Error Terms 1–18; RI = real–ideal discrepancy; RO = real–ought discrepancy; DEP = depression; ANX = anxiety; a = Time 1; b = Time 2; PC = Personal Constructs; CC = Conventional Constructs; AM = Abstract Measures; BDI = Beck Depression Inventory–II, CES-D = Center for Epidemiologic Studies–Depression Scale; DASS-D = Depression Anxiety Stress Scales–Depression scale; BAI = Beck Anxiety Inventory; STAI-T-A = State–Trait Anxiety Inventory–Trait scale–Anxiety factor; DASS-A = DASS–Anxiety scale; p1–p3 = Parcels 1–3.

measurement properties across a 4-week period. Test–retest reliability correlations over 4 weeks were strong for latent variables RI (.92) and RO (.91), and they were above Joiner et al.'s (2005) benchmark of .70 for PC-RI (.72) and PC-RO (.74), for CC-RI (.77) and CC-RO (.72), but not for AM-RI (.64) and AM-RO (.59). However, the test–retest coefficients for CC-RI and CC-RO were inflated by the stability of their error terms. RI and RO latent variables were factorially distinct. Convergent evidence of validity was moderate to strong for all measures except AM-RO. The PC and CC measures showed stronger convergence than did the AM at both time points, and the PC measures showed stronger convergence than did the CC measures at Time 1. Discriminant evidence of validity was found for all individual measures except AM-RO. Predictive validity of test-criterion relationships over 4 weeks, with three clinical measures each of anxiety and depression as criteria, was supported for latent variables RI and RO, for PC-RI and PC-RO, and for CC-RI; however, these relationships were not as strong for CC-RO and AM-RI, and AM-RO failed to predict two measures of anxiety. In summary, comparisons of the psychometric properties of the three instruments showed that the properties of the idiographic PC measures were consistently as strong as or stronger than those of the CC measures and the AM, as also was found for the clinical sample.

Study 2B

In Study 2B we examined measurement invariance across the two samples to determine whether the instruments' scores have comparable measurement properties in clinical and nonclinical populations. We also sought to strengthen the test-criterion evidence of validity of the discrepancy variables by conducting a known-groups analysis in which the clinical and nonclinical samples were compared. We expected the clinical sample to have higher mean scores on the discrepancy latent variables, as well as on the anxiety and depression latent variables. Whereas the test-criterion evidence of validity in Studies 1 and 2A relied only on self-reported criterion variables (i.e., self-reported anxiety and depression), this study also made use of an objective criterion (i.e., group membership).

Results

For the nonclinical sample, we modeled the discrepancy variables, using Time 1 data, so that all discrepancy data would be participants' first reports. The anxiety variable was modeled with three random parcels of STAI-T-A items, and the depression variable was modeled with three random parcels of BDI-II items, so that the indicators of the distress variables would be the same as those for the clinical sample. All parcels had normal skew and kurtosis except for one BDI-II parcel, which had abnormal kurtosis. The anxiety and depression parcels were left in their raw metric because maximum likelihood estimation is robust against mild violations of normality.

Equivalence of factor loadings and intercepts across groups (measurement invariance). Fit indexes for models tested in this study are reported in Table 6. Measurement invariance across groups can be tested in a way that is directly analogous to testing measurement invariance across time (Widaman & Reise, 1997); measurement invariance ensures construct comparability across groups. We conducted a two-group, four-factor means-and-covariance-structures analysis of the RI, RO, anxiety, and depression latent variables. The baseline model was identical in structure to Model 3 from Study 1 (Table 3), except that a means-and-intercepts structure was added and the model was conducted as a two-group analysis. The model was identified by standardizing the latent variables in the nonclinical sample and constraining the loadings and intercepts of the PC observed variables to be equal across groups. This model had good fit (Table 6, Model 1). Constraining loadings and intercepts to be equal across groups resulted in no loss of fit ($\text{TLI}_{\text{diff}} = .00$; Table 6, Models 2 and 3). These findings establish the comparability across groups of the RI, RO, anxiety, and depression latent variables and their relations to scores from the individual instruments.

Criterion relationships of latent variables. To test for group differences in anxiety and depression and to evaluate the criterion relationships of RI and RO, the means of the latent variables were constrained to be equal across groups. All four constraints led to a significant decline in fit, $\chi^2_{\text{diff}}(1, Ns = 278, 125) \geq 32.32$, $ps < .001$ (Table 6, Models 4–7), indicating that the latent variable means of the clinical sample (RI, $M = .80$; RO, $M = .91$; depression, $M = 1.36$; anxiety, $M = .95$) were higher than those of the nonclinical sample, each of which had been set to equal zero.

In summary, the results of Study 2B showed measurement invariance across the two samples, indicating that the instruments' scores have comparable measurement properties in both clinical and nonclinical populations (e.g., that the same difference in scores can be interpreted in the same way in both populations). The results also support the criterion relationships of the RI and RO

Table 6
Model Goodness-of-Fit Indicators for Study 2B (Ns = 125 and 278)

Model	*df*	χ^2	TLI	CFI	RMSEA
Model 1: Baseline MACS model	90	126.36**	.98	.99	.03
Model 2: Invariance of loadings	98	140.03**	.98	.99	.03
Model 3: Invariance of intercepts	106	157.66**	.98	.98	.04
Model 4: Equality of RI latent variable means	107	189.98***	.97	.97	.04
Model 5: Equality of RO latent variable means	107	206.63***	.96	.97	.05
Model 6: Equality of depression latent variable means	107	237.57***	.95	.96	.06
Model 7: Equality of anxiety latent variable means	107	207.73***	.96	.97	.05

Note. TLI = Tucker–Lewis index; CFI = comparative fit index; RMSEA = root-mean-square error of approximation. MACS = means-and-covariance structures; RI = real–ideal discrepancy; RO = real–ought discrepancy.
** $p < .01$. *** $p < .001$.

latent variables by showing that their means, as well as the means of the anxiety and depression latent variables, were higher in the clinical sample than in the nonclinical sample.

General Discussion

In the present study, using a clinical sample and a nonclinical sample, we evaluated and compared the psychometric properties of three instruments that assess the RI and the RO self-discrepancies: the idiographic SCQ-PC, the nonidiographic SCQ-CC, and the content-free AM.

In preliminary analyses, the criticism that self-discrepancies measured with difference scores have weak reliability and validity (e.g., Cronbach & Furby, 1970; Wylie, 1974) was contradicted. Results showed that the RI and RO difference scores used in the PC and CC instruments have test–retest reliability coefficients that are at least as strong as those of the real-self, ideal-self, and ought-self components of the discrepancies. Also, results of multiple regression analyses supported the hypothesis that the difference scores generally have incremental validity of test-criterion relationships to anxiety and depression scores but that the real-self score advocated by Wylie (1974), as well as the ideal-self and ought-self scores, do not, suggesting that the distance between the components conveys important information. These findings support the reliability and validity of difference scores in measuring the phenomenological construct self-discrepancy, as did similar findings with an early version of the PC-RI measure (Watson & Watts, 2001).

CFA provided an exacting evaluation and comparison of the test–retest reliabilities and of convergent and discriminant evidence of the validities of the RI and RO measures in the three self-discrepancy instruments, as well as an evaluation of the properties of the RI and RO latent variables. The correlated uniqueness CFA model we used rests on the assumption that the variance unique to a particular observed-variable indicator of the latent variable is extraneous to the underlying construct, which is consistent with the present evaluation of instruments designed to measure the same latent variables. The latent variables' strong test–retest correlations (.92 for RI and .91 for RO) and significant criterion relationships to the anxiety and depression latent variables (all *p*s $< .001$) provided a sound basis for the evaluation of the observed variables. Also, results showed measurement invariance across time and across the two samples, the latter indicating that the instruments' scores have comparable measurement properties in both clinical and nonclinical populations (e.g., that the same difference in scores can be interpreted in the same way in both populations).

Internal consistency alpha coefficients were strong for PC-RI and PC-RO (.90 to .92) and were moderate to strong for CC-RI and CC-RO (.82 to .90) across both samples. Average interitem correlations all fell within Clark and Watson's (1995) ideal range of .15 to .50. Internal consistency is not applicable to the one-item AM. In the clinical sample, the stronger PC alphas in comparison to CC alphas suggest, as would be expected, that more of the items in the idiographic instrument reflect discrepancy than do the items in the nonidiographic instrument.

Test–retest reliability coefficients over 4 weeks were above Joiner et al.'s (2005) benchmark of .70 for PC-RI (.72) and PC-RO (.74), for CC-RI (.77) and CC-RO (.72), but not for AM-RI (.64) and AM-RO (.59). The coefficients for the PC and CC measures were higher than those reported for the Selves Questionnaire (Higgins, 1987; Moretti & Higgins, 1990; Scott & O'Hara, 1993). However, CFA findings showed that the observed correlations for CC-RI and CC-RO and for AM-RI and AM-RO are inflated by stability of their error. The findings provide the strongest support for the test–retest reliability of the PC measures.

Discriminant evidence of validity was found for the RI and RO latent variables in both samples, indicating that RI and RO are distinct constructs. Regarding the ability of the individual instruments to distinctively measure these constructs, convergent evidence and discriminant evidence of validity were found for all the measures of RI and RO in both samples, except for AM-RO in the nonclinical sample. Discriminant evidence of validity was found even though the RI and RO latent variables were strongly correlated and the RI and RO observed variables were moderately to strongly correlated. Convergent evidence of validity was as strong or stronger for the PC measures compared with the CC measures, and convergent evidence of validity was stronger for the PC measures and generally for the CC measures than it was for the AM.

Test-criterion evidence of validity, using multiple clinical measures of anxiety and depression as criteria, was found for the RI and RO latent variables and for all RI and RO individual measures except for AM-RO in relation to anxiety. In the clinical sample, we found these test-criterion relationships at one time point, using one measure each of anxiety and depression. In the nonclinical sample, we found predictive validity across 4 weeks, using three measures each of anxiety and depression. However, CFA findings showed that the test-criterion relationships for CC-RO, AM-RI, and AM-RO were compromised by the correlations of their error terms with the criteria. Thus, the findings provide the strongest support for the test-criterion relationships of the PC instrument. Previous research also has found stronger test-criterion relationships for an idiographic self-discrepancy measure compared with a nonidiographic one (Moretti & Higgins, 1990; Watson & Watts, 2001). Test-criterion evidence of validity was also shown by the higher means of the anxiety and depression latent variables and the RI and RO latent variables in the clinical sample compared with the nonclinical sample.

In summary, the findings in both the clinical and nonclinical samples support the psychometric properties of the PC instrument most strongly and the properties of the CC instrument less strongly for internal consistency, test–retest reliability, and convergent evidence of validity. The properties of the AM instrument are the weakest; however, AM-RI is stronger than AM-RO in discriminant evidence of validity and test-criterion evidence of validity. Even though AM-RI, which is a one-item measure, has weaker test–retest reliability and convergent evidence of validity than do the PC-RI and CC-RI, it may be useful because of its fast administration.

A limitation of the present study is that clinical interviews were not included in the assessment of anxiety and depression, although each participant was assessed with three instruments. Also, generalization of the results is limited by the primarily White American university student samples.

The findings have implications for self-discrepancy theory. The CFA test of discriminant evidence of validity, which to our knowledge is the first such test, showed that RI and RO are distinct discrepancies. This finding is consistent with Higgins's (1987) assumption that there are two discrepancies and contradicts suggestions that only a single discrepancy exists (e.g., Phillips & Silvia, 2005; Tangney et al., 1998). The finding that the two discrepancies are strongly correlated, as others also have found (e.g., Phillips & Silvia, 2005; Tangney et al., 1998), is consistent with Rogers's (1959) theory that a high RO leads to a high RI (see the introductory section).

A clinical implication of the findings of the present study and of past studies (e.g., Tangney et al., 1998; Weilage & Hope, 1999) is that a negative self-evaluation for not meeting one's own expectations and for not meeting one's perception of others' expectations is experienced in anxiety and depression. A therapeutic focus on the experience of self-discrepancy may contribute to a positive outcome in therapy. Rogers (1959) theorized that self-discrepancy is a personality predisposition to emotional distress and that empathic communication by the therapist decreases self-discrepancy and thereby decreases anxiety and depression. The relationship of therapist empathy to therapy outcome has been found in studies of various schools of therapy (Bohart, Elliott, Greenberg, & Watson, 2002). The idiographic PC instrument assesses self-discrepancy in the individual's own words, which the therapist can use to increase empathic communication.

Of the three instruments, the findings most strongly support the psychometric properties of the idiographic PC instrument for use in clinical practice and in research. If a research protocol can accommodate all three instruments, the stronger test–retest correlations of the latent variables compared with those of the observed variables make statistical tests more powerful than they would be with only the PC instrument. Areas of research include personality and clinical studies of Rogers's (1959) and Higgins's (1987) theories of self-discrepancy as a personality predisposition to emotional distress and clinical studies of psychotherapy process and outcome.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Arbuckle, J. L. (2006). Amos 7.0 [Computer software]. Spring House, PA: Amos Development Corp.

Barrett-Lennard, G. T. (1962). Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs, 76*(43, Whole No. 562).

Beck, A. T., & Steer, R. A. (1990). *Manual for the Beck Anxiety Inventory.* San Antonio, TX: Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory* (2nd ed.). San Antonio, TX: Psychological Corporation.

Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 3–20). Madison, WI: University of Wisconsin Press.

Bieling, P. J., Antony, M. M., & Swinson, R. P. (1998). The State–Trait Anxiety Inventory, Trait Version: Structure and content re-examined. *Behaviour Research and Therapy, 36,* 777–788.

Bohart, A. C., Elliott, R., Greenberg, L. S., & Watson, J. C. (2002). Empathy. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 89–108). New York, NY: Oxford University Press.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York, NY: Wiley.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York, NY: Guilford.

Bruch, M. A., Rivet, K. M., & Laurenti, H. J. (2000). Type of self-discrepancy and relationships to components of the tripartite model of emotional distress. *Personality and Individual Differences, 29,* 37–44.

Bryan, B. C., Watson, N., Babel, K. S., & Thrash, T. M. (2008, June). *Relations of self-discrepancies to anxiety and depression: Tests of Rogers's and Higgins's theories.* Paper presented at the annual meeting of the Society for Psychotherapy Research, Barcelona, Spain.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7,* 309–319.

Cronbach, L. J., & Furby, L. (1970). How we should measure change—Or should we? *Psychological Bulletin, 74,* 68–80.

Fairbrother, N., & Moretti, M. (1998). Sociotropy, autonomy, and self-discrepancy: Status in depressed, remitted depressed, and control participants. *Cognitive Therapy and Research, 22,* 279–297.

Francis, J. J., Boldero, J. M., & Sambell, N. L. (2006). Self-lines: A new, psychometrically sound, "user-friendly" idiographic technique for assessing self-discrepancies. *Cognitive Therapy and Research, 30,* 69–84.

Gough, H. G., & Heilbrun, A. B., Jr. (1965). *Adjective Check List manual.* Palo Alto, CA: Consulting Psychologists Press.

Gramzow, R. H., Sedikides, C., Panter, A. T., & Insko, C. A. (2000). Aspects of self-regulation and self-structure as predictors of perceived emotional distress. *Personality and Social Psychology Bulletin, 26,* 188–205.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94,* 319–340.

Higgins, E. T., Klein, R., & Strauman, T. (1985). Self-concept discrepancy theory: A psychological model for distinguishing among different aspects of depression and anxiety. *Social Cognition, 3,* 51–76.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

James, W. (1981). The consciousness of self. In W. James (Ed.), *The principles of psychology* (Vol. 1, pp. 279–379). Cambridge, MA: Harvard University Press. (Original work published 1890)

Joiner, T. E., Jr., Walker, R. L., Pettit, J. W., Perez, M., & Cukrowicz, K. C. (2005). Evidence-based assessment of depression in adults. *Psychological Assessment, 17,* 267–277.

Kelly, G. A. (1955). *The psychology of personal constructs* (Vols. 1–2). New York, NY: Norton.

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait–multimethod matrix by confirmatory factor analysis. *Psychological Bulletin, 112,* 165–172.

Kinderman, P., & Bentall, R. (1996). Self-discrepancies and persecutory delusions: Evidence for a model of paranoid ideation. *Journal of Abnormal Psychology, 105,* 106–113.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Linn, R. L., & Slinde, J. A. (1977). The determination of the significance of change between pre- and posttesting periods. *Review of Educational Research, 47,* 121–150.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32,* 53–76.

Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 21–38). Madison, WI: University of Wisconsin Press.

Lovibond, S. H., & Lovibond, P. F. (1993). *Manual for the Depression*

Anxiety Stress Scales (DASS) (Psychology Foundation Monograph. Available from Psychology Foundation, Room 1005 Mathews Building, University of New South Wales, NSW 2052, Australia).

Marsh, H. W. (1989). Confirmatory factor analysis of multitrait–multimethod data: Many problems and a few solutions. *Applied Psychological Measurement, 13,* 335–361.

McDonald, R. P., & Ho, M.-H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7,* 64–82.

Meng, X., Rosenthal, R., & Rubin, R. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111,* 172–175.

Meyer, A. E. (1981). The Hamburg short psychotherapy comparison experiment. *Psychotherapy and Psychosomatics, 35,* 81–207.

Moretti, M. M., & Higgins, E. T. (1990). Relating self-discrepancy to self-esteem: The contribution of discrepancy beyond actual-self ratings. *Journal of Experimental Social Psychology, 26,* 108–123.

Parker, G. V., & Veldman, D. J. (1969). Item factor structure of the Adjective Check List. *Educational and Psychological Measurement, 29,* 609–613.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609.

Phillips, A. G., & Silvia, P. J. (2005). Self-awareness and the emotional consequences of self-discrepancies. *Personality and Social Psychology Bulletin, 31,* 703–713.

Pitts, S. C., West, S. G., & Tein, J.-Y. (1996). Longitudinal measurement models in evaluation research: Examining stability and change. *Evaluation and Program Planning, 19,* 333–350.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88,* 879–903.

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.

Rogers, C. R. (1951). *Client-centered therapy: Its current practice, applications and theory.* Boston, MA: Houghton Mifflin.

Rogers, C. R. (1959). A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of a science: Vol. 3. Formulations of the person and the social context* (pp. 184–256). New York, NY: McGraw-Hill.

Rogers, C. R., & Dymond, R. F. (Eds.). (1954). *Psychotherapy and personality change: Coordinated research studies in the client-centered approach.* Chicago, IL: University of Chicago Press.

Scott, L., & O'Hara, M. W. (1993). Self-discrepancies in clinically anxious and depressed university students. *Journal of Abnormal Psychology, 102,* 282–287.

Shlien, J. M. (1962). Toward what level of abstraction in criteria? In H. H. Strupp & L. Luborsky (Eds.), *Research in psychotherapy* (Vol. 2, pp. 142–154). Washington, DC: American Psychological Association.

Spielberger, C. D. (1983). *Manual for the State–Trait Anxiety Inventory (Form Y).* Palo Alto, CA: Consulting Psychologists Press.

Strauman, T. J. (1989). Self-discrepancies in clinical depression and social phobia: Cognitive structures that underlie emotional disorders? *Journal of Abnormal Psychology, 98,* 14–22.

Strauman, T. J. (1992). Self-guides, autobiographical memory, and anxiety and dysphoria: Toward a cognitive model of vulnerability to emotional distress. *Journal of Abnormal Psychology, 101,* 87–95.

Strauman, T. J. (1996). Stability within the self: A longitudinal study of the structural implications of self-discrepancy theory. *Journal of Personality and Social Psychology, 71,* 1142–1153.

Strauman, T. J., & Higgins, E. T. (1988). Self-discrepancies as predictors of vulnerability to distinct syndromes of chronic emotional distress. *Journal of Personality, 56,* 685–707.

Strauman, T. J., Kolden, G. G., Stromquist, V., Kwapil, L., Schneider, K., Heerey, E., & Schneider, K. (2001). The effects of treatments for depression on perceived failure in self-regulation. *Cognitive Therapy and Research, 25,* 693–712.

Strauman, T. J., Vookles, J., Berenstein, V., Chaiken, S., & Higgins, E. T. (1991). Self-discrepancies and vulnerability to body dissatisfaction and disordered eating. *Journal of Personality and Social Psychology, 61,* 946–956.

Tangney, J. P., Niedenthal, P. M., Covert, M. V., & Barlow, D. H. (1998). Are shame and guilt related to distinct self-discrepancies? A test of Higgins's (1987) hypotheses. *Journal of Personality and Social Psychology, 75,* 256–268.

Tashiro, T., & Mortensen, L. (2006). Translational research: How social psychology can improve psychotherapy. *American Psychologist, 61,* 959–966.

Vautier, S., Callahan, S., Moncany, D., & Sztulman, H. (2004). A bistable view of single constructs measured using balanced questionnaires: Application to trait anxiety. *Structural Equation Modeling, 11,* 261–271.

Watson, N. (2004). *Self-Concept Questionnaire–Personal Constructs, Self-Concept Questionnaire—Conventional Constructs, Abstract Measures of Real-Ideal and Real-Ought Discrepancies.* Retrieved from http://www.wm.edu/research/watson

Watson, N., & Watts, R. H., Jr. (2001). The predictive strength of personal constructs versus conventional constructs: Self-image disparity and neuroticism. *Journal of Personality, 69,* 121–143.

Weilage, M., & Hope, D. A. (1999). Self-discrepancy in social phobia and dysthymia. *Cognitive Therapy and Research, 23,* 637–650.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). Washington, DC: American Psychological Association.

Wylie, R. C. (1974). *The self-concept: Vol. 1. A review of methodological considerations and measuring instruments* (rev. ed.). Lincoln, NE: University of Nebraska Press.

Received March 30, 2009
Revision received April 20, 2010
Accepted June 11, 2010 ■