Personal RELATIONSHIPS
JOURNAL OF THE INTERNATIONAL ASSOCIATION FOR RELATIONSHIP RESEARCH

Personal Relationships, (2012). Printed in the United States of America.
 DOI: 10.1111/j.1475-6811.2012.01412.x

The positive and negative semantic dimensions of relationship satisfaction

RICHARD E. MATTSON,[a] RONALD D. ROGGE,[b] MATTHEW D. JOHNSON,[c] ELIZABETH K. B. DAVIDSON,[b] AND FRANK D. FINCHAM[d]

[a]*Auburn University;* [b]*University of Rochester;* [c]*State University of New York at Binghamton;* [d]*Florida State University*

Abstract

Semantic differential items were reconfigured to assess relationship satisfaction across separate positive and negative attitude dimensions. Study 1 ($N = 1,656$) supported a 2-factor model for the Positive and Negative Semantic Differential (PN–SMD), as well as its convergent, criterion-related, and incremental validity over the 16-item Couples Satisfaction Index (CSI; J. L. Funk & R. D. Rogge, 2007) using known correlates of relationship satisfaction as criteria. Study 2 ($N = 89$) replicated the convergent, criterion-related, and incremental validity findings of Study 1 using different criterion measures, the CSI, a bipolar semantic differential measure designed for assessing relationship satisfaction, and an existing 2-dimensional measure of relationship satisfaction. The authors demonstrated across studies that the PN–SMD captures criterion-relevant information about ambivalence versus indifference toward the relationship—associations that are only detectable when using a 2-dimensional satisfaction measure.

Relationship satisfaction refers to a person's overall evaluation of his or her relationship. It is arguably the most important variable in relationship research. Indeed, indices of relationship satisfaction are the most common metric for assessing how and in what ways different aspects of a relationship relate to its overall functioning (Jacobson, 1985). It is also the de facto criterion used to validate developmental models of relationship discord, as well as the gold standard for evaluating interventions designed for its alleviation.

Indices of relationship satisfaction typically place favorable versus unfavorable judgments about one's relationship at opposing ends of a single attitude dimension. For example, several one-dimensional measures of relationship satisfaction ask participants to rate their overall relationship on scales ranging from *extremely unhappy* to *perfect*. Because these attitudes are placed on the same response continuum, endorsing favorable evaluations toward the relationship (e.g., *very happy*) equates to the rejection of unfavorable ones (i.e., *a little unhappy*). This effectively reduces the definition of relationship satisfaction to the relative absence of dissatisfaction and vice versa (Bradbury, Fincham, & Beach, 2000; also see Cacioppo, Gardner, & Bernston, 1997). Despite the pervasive application of this formulation in measurement, many consider it to be an inadequate model of relationship satisfaction for several reasons (e.g., Weiss, 2005).

Foremost, the constructs of satisfaction and dissatisfaction are not polar opposites. Indeed, being completely satisfied with one's relationship likely means something more than just the absence of negative evaluations. Rather, being completely satisfied (i.e., holding maximally favorable evaluations toward the relationship) is, more accurately, the opposite of being completely *unsatisfied*—that is, holding no favorable evaluations

Richard E. Mattson, Department of Psychology, Auburn University; Ronald D. Rogge, Department of Clinical and Social Sciences in Psychology, University of Rochester; Matthew D. Johnson, Department of Psychology, State University of New York at Binghamton; Elizabeth K. B. Davidson, Department of Clinical and Social Sciences in Psychology, University of Rochester; Frank D. Fincham, Family Institute, Florida State University.

Correspondence should be addressed to Richard E. Mattson, Department of Psychology, Auburn University, 226 Thach Hall, Auburn, AL 36849, e-mail: rem0003@auburn.edu

about the relationship. Likewise, being completely unsatisfied is not synonymous with being completely *dissatisfied*, as having few favorable evaluations toward a relationship is different from evaluating it as maximally bad. Also, as most relationships comprise both satisfying and dissatisfying characteristics (e.g., Guilford & Bengtson, 1979; Orden & Bradburn, 1968), individuals likely hold both favorable and unfavorable evaluations about their relationship simultaneously. Placing these attitudes on the same bipolar response scale, therefore, creates interpretive difficulties. For example, responses around a scale's midpoint may alternatively represent *indifference* (i.e., neither highly favorable nor unfavorable evaluations) or *ambivalence* (i.e., strongly favorable and unfavorable evaluations held simultaneously; see Klopfer & Madden, 1980). Taken together, one-dimensional measures may inaccurately model overall attitudes toward the relationship and, as a consequence, may "yield ambiguous findings and contribute little to an understanding of [relationship] process" (Johnson, White, Edwards, & Booth, 1986, p. 42).

A two-dimensional model of relationship satisfaction

In light of the abovementioned concerns, Fincham, Beach, and Kemp-Fincham (1997) and Fincham and Linfield (1997) proposed a two-dimensional model of relationship satisfaction comprising separate evaluations of the positive (i.e., satisfying) and negative (i.e., dissatisfying) aspects of the relationship. They developed a brief six-item measure entitled the Positive and Negative Quality in Marriage Scale (PANQIMS). The constituent positive and negative subscales each contain three items that differ only with respect to the dimension of relationship satisfaction they measure (e.g., "Considering the positive qualities of your spouse, and ignoring the negative qualities, evaluate how positive these qualities are"). Using confirmatory factor analysis, they found that the PANQIMS items conformed best to an oblique two-factor model versus a solution in which all items loaded onto a single dimension. In other words, the better fitting model treated the evaluation of positive and negative relationship characteristics as two distinct, albeit correlated, attitudes. Also, using multiple regression analysis, the separate positive and negative dimensions predicted unique variance in self-reports of dyadic behavior and relationship attributions beyond the Marital Adjustment Test (MAT; Locke & Wallace, 1959), a popular one-dimensional measure of relationship satisfaction. This suggests that positive and negative attitudes toward the relationship have distinct associations with important aspects of relationship functioning—that the information gained in separately assessing positive and negative attitudes has added value.

Also important, Fincham and Linfield (1997) found evidence that separately measuring the positive and negative dimensions of relationship satisfaction can highlight important differences in attitudes of ambivalence versus indifference toward the relationship. They defined ambivalence as being high on both dimensions (i.e., holding both favorable and unfavorable attitudes), whereas indifference was characterized as being low in both positive and negative attitudes (i.e., relatively neutral sentiments about the relationship). They found that when compared to individuals reporting indifference, individuals identified as ambivalent reported higher ratios of negative to positive dyadic behaviors and endorsed more maladaptive attributions regarding their partners' behavior. Importantly, however, ambivalent and indifferent individuals did not differ with respect to their MAT scores. That is, the qualitatively different attitudes held by indifferent and ambivalent individuals were not distinguishable on the basis of a one-dimensional measure of relationship satisfaction.

Taken together, Fincham and Linfield's (1997) findings provide empirical support that (a) evaluations of the positive and negative aspects of the relationship are separable attitudes and (b) constraining them to a one-dimensional scale can lose important information about a relationship's evaluation. Mattson, Paldino, and Johnson (2007) replicated these findings using the semantic differential as a one-dimensional measure

of relationship satisfaction and, as criteria, maladaptive attributions and observations of conflict behaviors (also see Abakoumkin, Stroebe, & Stroebe, 2010; Menchaca & Dehle, 2005).

Measurement limitations of the PANQIMS

Findings with the PANQIMS support a two-dimensional model of relationship satisfaction, as well as its incremental validity over validated one-dimensional scales. However, Fincham and Linfield (1997) did not explicitly define the operational domain of positive and negative relationship satisfaction, so it is difficult to evaluate whether items on the PANQIMS represent the best way to measure these constructs.

The PANQIMS also has at least two notable limitations. First, items within each of the PANQIMS subscales assess relationship satisfaction using a single type of descriptive continuum (e.g., "how *positive*"), leaving out other evaluative descriptors that may also compose these potentially multifaceted attitudes (e.g., "how *exciting, interesting*, etc."). Second, each of its three items per dimension vary with respect to either the object being evaluated (i.e., partner or relationship), or which aspect of the object is under evaluation (i.e., "qualities of" or "feelings toward" the relationship). As such, some of the item content (e.g., *feelings* toward a *partner's* qualities) may extend beyond the definition of relationship satisfaction as an attitude toward the relationship. If so, then empirical associations between the PANQIMS and other variables become difficult to interpret.

Toward a better two-dimensional measure

It is not surprising that Fincham and Linfield (1997) left unspecified the exact operational domain of positive and negative relationship satisfaction. Their focus was on neither construct development nor psychometric validation of the PANQIMS. Rather, it was to explore the potential utility of measuring relationship satisfaction across positive and negative dimensions. As such, a more thoroughgoing psychometric process was likely premature for such an exploratory endeavor. Furthermore, "scientists seldom outline the domain of observables before assuming that any one ... relates to a construct [and] ... should not be criticized for provisionally assuming that particular observables relate to a vaguely understood construct" (Nunnally & Bernstein, 1994, p. 87). Indeed, the process of construct validation is more iterative than linear and entails additional attempts to revise or redefine the domain, as well as a continued dialogue between theory and empirical research.

To that end, we sought to extend Fincham and Linfield's (1997) work by (a) specifying a theory-driven measurement model for relationship satisfaction and, from that, (b) constructing a new measure of positive and negative relationship satisfaction. Our measurement model derives from the semantic differential (Osgood, Suci, & Tannenbaum, 1957). The semantic differential, in brief, measures attitudes toward a particular object (e.g., an intimate relationship) using bipolar continua anchored by adjectives of opposing meaning (e.g., good–bad). There exists theoretical and empirical support for measuring relationship satisfaction via bipolar semantic differential items (e.g., Funk & Rogge, 2007; Karney & Bradbury, 1997). There is also evidence to support a two-dimensional model of relationship satisfaction. The proposed measure combines these two lines of research by separating semantic differential items into positive and negative subscales. We used validated means for doing so (see Kaplan, 1972), which are based on a two-dimensional model of attitudes. We proceed with a brief synopsis of the bipolar semantic differential and its psychometric foundation, and then incorporate the construct of relationship satisfaction into its measurement model. We then discuss our approach to item selection and scale construction, as well as our means for evaluating the measure's psychometric validity.

Conceptual and psychometric foundations

The measurement of meaning

Osgood and colleagues (1957) argued that the meaning of a particular object (concept,

activity, etc.) to an individual can be quantitatively measured. They construed *meaning* as a process that involves the appraisal of an object across common experiential dimensions. They contended that such appraisals can be measured using response scales with bipolar semantic endpoints (e.g., -3 = *dissatisfied*, $+3$ = *satisfied*). They justified this approach on the following grounds: Words (e.g., satisfaction) accrue meaning through repeated pairings with specific experiences (e.g., feeling satisfied) that, over time, give the word its semantic value. Objects (e.g., one's relationship) are also repeatedly paired with experiential outcomes, which lead to the formation of attitudes: a "learned predisposition to respond to [an object] in a consistently favorable or unfavorable manner" (Allport, 1935, p. 818; also see Olson & Fazio, 2001). Because particular words (e.g., satisfied) denote specific experiences (e.g., feeling satisfied), an individual's attitude toward any object (e.g., one's relationship) can be understood with reference to these semantic anchors (e.g., "How satisfying is your relationship?"). Osgood and colleagues conducted factor analyses on several items varying in semantic content. Their data revealed three primary axes underlying meaning (i.e., attitudes), which were interpreted to represent the common experiential dimensions of evaluation (e.g., good–bad), potency (e.g., strong–weak), and activity (e.g., fast–slow). They argued that by using bipolar semantic continua that represent these different experiential dimensions, one can quantifiably measure the strength and valence of an attitude toward any given object.

The semantic differential as a measurement model for relationship satisfaction

Measures of relationship satisfaction frequently lack a clear theoretical framework (with notable exceptions; e.g., see Rusbult, Martz, & Agnew, 1998), often relying on implicit theories about relationship functioning to guide item selection. In contrast, semantic differential items derive from a theoretical basis that is highly compatible with contemporary formulations of relationship satisfaction and its determinants. For instance, relationship satisfaction is thought to comprise an individual's judgment about his or her relationship and, thus, can be considered a particular case within the larger domain of constructs measured by the semantic differential (i.e., attitudes; see Bradbury et al., 2000). Moreover, the semantic differential assumes that the number of pairings between experientially aversive or appetitive outcomes and a particular object determines its subjective appraisal. This model fits squarely into developmental theories of relationship satisfaction, which typically assume that the trajectory of favorable versus unfavorable evaluations result from the exchange of rewarding or aversive relationship outcomes, respectively (see Karney & Bradbury, 1995). Taken together, the semantic differential provides a theoretical link between the conceptualization of relationship satisfaction and its measurement, and between variability in relationship functioning and observable differences in responses to evaluative semantic continua (for similar arguments, see Fincham & Bradbury, 1987).

The semantic differential as a global measure of relationship satisfaction

The semantic differential measures an individual's overall (or global) attitude toward a particular object, a focus that is highly consistent with the "pervasive tendency on the part of marital researchers to favor global evaluations ... in the measurement of [satisfaction toward the relationship]" (Fincham & Bradbury, 1987, p. 799). However, there is another reason why it is important to constrain measures of relationship satisfaction to attitudinal judgments about the overall relationship. Specifically, it decreases the risk of incorporating content pertaining to other, operationally distinct constructs, such as communication, love, trust, and so on (for a similar argument, see Norton, 1983). This differentiates global measures of relationship satisfaction from omnibus ones, such as the Dyadic Adjustment Scale (DAS; Spanier, 1976). For example, the DAS includes items that assess the frequency with which partners "leave the house after a fight" and

other items more representative of communication styles. Although there are advantages to omnibus scales—especially in clinical settings—global measures are especially useful in research examining the correlates and predictors of relationship satisfaction.

The work of Funk and Rogge (2007) supports this contention, as well as highlights the utility of semantic differential items in assessing relationship satisfaction. In particular, they evaluated several predominant relationship satisfaction measures using IRT analysis, a technique that can identify the precision of items and the information they provide about a latent construct (e.g., relationship satisfaction). They found that items more descriptive of relationship functioning (e.g., agree on recreation) were relatively indistinct and uninformative with respect to relationship satisfaction. Indeed, many of these items loaded more on a dyadic communication factor, as opposed to the construct of relationship satisfaction in exploratory factor analyses. By comparison, semantic differential items were highly precise and provided some of the most information about an individual's attitude toward his or her relationship. Considered together, the semantic differential items are an efficient means to assess global attitudes toward the relationship and do so without contaminating this evaluative judgment with reports on its hypothesized determinants (e.g., communication patterns).

Item selection and test construction

Although serving as the basis of our measurement model, the initial semantic differential provides a rather large pool of candidate items, some of which may be less applicable to appraising relationships (e.g., angular–rounded) than are others (e.g., interesting–boring). As such, we restricted our options to the semantic differential measure developed by Karney and Bradbury (1997). This measure was constructed specifically for assessing relationship satisfaction and has been validated for this purpose. From this pool, we selected the seven items that Funk and Rogge (2007) identified as the most precise and informative with respect to the latent construct of relationship satisfaction, allowing for a more efficient (i.e., less redundant) measure.

This item set has three other appealing characteristics. First, while the items primarily tap the *evaluative* axis of Osgood and colleagues (1957) three-dimensional taxonomy; some correlate to varying degrees with the dimensions of potency (e.g., *sturdy–fragile*) and activity (e.g., *interesting–boring*). As such, these semantic differential items may access other attitude dimensions proven relevant to overall evaluations of one's relationship, although still maintaining consistency with the conceptualization of relationship satisfaction as a single, global, and primarily evaluative judgment. Second, there already exists a validated method for creating positive and negative subscales from bipolar semantic differential items (see Kaplan, 1972). Specifically, the semantic anchors of each bipolar adjective set (e.g., *good–bad*) are uncoupled and placed on their own response continuums. By contrast, selecting a different measure (e.g., the quality in marriage index [QMI]; Norton, 1983) could preclude use of this method, as separating certain items into positive and negative components may be difficult (e.g., "I really feel like part of a team with my partner"). Third, some of the items designed by Karney and Bradbury (1997) comprise adjective pairings that are not exact polar opposites (e.g., *hopeful–discouraging*, as opposed to *hopeful–hopeless*), a feature that is particularly advantageous in the current context. Specifically, individuals may be less inclined to treat the semantic anchors as bipolar, thus reducing response consistency or carryover effects, respectively, across items and subscales (for a similar approach, see Cacioppo et al., 1997).

Integrative summary

Evidence suggests that the two-dimensional focus of the PANQIMS more accurately represents evaluations of the relationship. However, there is reason to believe that the semantic differential may improve the measurement of positive and negative relationship satisfaction. Specifically, the semantic

differential affords a wider range of potential item content than the PANQIMS, while also restricting the operational definition to attitudes toward the overall relationship. Also, along with the semantic differential comes a theoretical model for attitude formation that is consistent with predominant formulations of relationship satisfaction and its hypothesized determinants (see Bradbury et al., 2000; Karney & Bradbury, 1995). The semantic differential items also have several appealing characteristics relative to other potential candidates. For example, the semantic differential items outperform those from many other, similarly global measures of relationship satisfaction (e.g., the QMI). Likewise, the use of semantic differential items may prevent the assessment of relationship satisfaction's hypothesized determinants—such as dyadic communication or compatibility—from confounding any observed associations between them. Taken together, research and theory support a two-dimensional conceptualization of relationship satisfaction, as well as the utility of semantic differential items in measuring this construct. Thus, combining these approaches—using a straightforward and valid means for separating semantic differential items into positive and negative indices—may improve the assessment of relationship satisfaction beyond existing measures of this construct. We evaluated this possibility across two studies.

Study 1

Using Fincham and Linfield's (1997) study as our guiding framework, we posed five primary hypotheses in Study 1. First, we predicted that the Positive and Negative Semantic Differential (PN–SMD) would conform to an oblique two-factor solution comprising a positive semantic dimension (PSD) and a negative semantic dimension (NSD), which would provide an incrementally better fit than an orthogonal two-factor or a one-factor solution. In other words, we hypothesized that Fincham and Linfield's CFA results using the PANQIMS would replicate using the PN–SMD. We also tested for measurement invariance across gender and relationship type (e.g., married, engaged, etc.)

Second, we predicted that elevated PSD and NSD scores would positively and negatively correlate, respectively, with reports of higher relationship satisfaction (convergent validity), as measured by the Couples Satisfaction Index (CSI; Funk & Rogge, 2007). This was done to establish that the PN–SMD subscales were, to some extent, measuring the same evaluative judgments as a validated one-dimensional measure with positive and negative endpoints.

Third, we hypothesized that the PN–SMD would demonstrate criterion-related validity using variables known to be strong correlates of relationship satisfaction (see Frame, Mattson, & Johnson, 2009). Specifically, we predicted that greater social support and sexual satisfaction would positively correlate with the PSD and negatively correlate with the NSD, whereas we expected the reverse pattern when using conflict or negative affectivity as the criterion. We also tested whether the correlations of the PN–SMD subscales differed significantly for each criterion. We anticipated that the PSD would more strongly correlate with positive aspects of the relationship (e.g., sexual satisfaction), whereas negative characteristics (e.g., conflict) would more strongly associate with the NSD. Significant findings demonstrate that the separate positive and negative attitudes toward the relationship carry meaningfully different associations in line with their putative determinants (i.e., the positive and negative characteristics of the relationship, respectively).

Fourth, we hypothesized that each PN–SMD subscale would account for unique variability in our criterion variables (i.e., incremental validity) beyond that which is predicted by the other semantic dimension and the CSI. In the former case, significant effects would indicate that aspects of relationship functioning uniquely correspond to separate positive and negative evaluative judgments when controlling for potential overlap. In the latter case, incremental validity would show that the information afforded by a two-dimensional scale is significantly more

than what is on offer from a one-dimensional index. Consistent with Fincham and Linfield's (1997) approach, we also tested whether the above associations were robust when controlling for the more general constructs of positive and negative affectivity.

Fifth, using the PN–SMD in lieu of PANQIMS, we sought to replicate Fincham and Linfield's (1997) finding that mean differences in the criteria would exist between satisfied (i.e., high PSD and low NSD), dissatisfied (i.e., low PSD and high NSD), ambivalent (i.e., high PSD and high NSD), and indifferent (i.e., low PSD and low NSD) participants. In addition to an omnibus test, we ran planned comparisons examining the ambivalent and indifferent groups. We predicted mean differences in the criteria across these two groups but also that ambivalent and indifferent participants would be indistinguishable on the basis of one-dimensional relationship satisfaction (i.e., CSI) scores.

We also explored whether these groups demonstrated dissimilar trajectories in one-dimensional relationship satisfaction. Of most interest was the comparison between the ambivalent and indifferent groups. Despite being indistinguishable at the cross-sectional level, we reasoned that ambivalent individuals may display steeper declines in satisfaction over time. Relative to indifferent participants, ambivalent individuals (by definition) have more highly favorable evaluations of their relationships' positive characteristics. However, as shown by Huston, Caughlin, Houts, Smith, and George (2001), a relationship's positive characteristics are more susceptible to decline over time, which, in turn, corresponds to decreases in global satisfaction. As such, it is possible that ambivalent individuals will be more prone to declines in global satisfaction over time. Demonstrating this effect would further challenge a one-dimensional operation of relationship satisfaction. Specifically, although ambivalent and indifferent individuals may at one time appear similar on a one-dimensional scale, underlying differences in their positive and negative relationship attitudes actually foreshadow alternate outcomes longitudinally.

Method

Participants

Respondents were at least 18 years old and were currently in a romantic relationship. Participants were recruited online (87%), from postings in online forums (e.g., The Knot.com and CanadianBride.com) and from the second author's website, as well as e-mails and e-mail distribution lists (8.2%), and in person with postcards (1.9%). The initial survey took approximately 20–25 min and participants were offered individualized feedback at the end of the survey. Via follow-up e-mails, we requested that participants complete the CSI (as well as other measures unrelated the current analyses) every 3 months for 18 months.

The initial sample comprised 1,727 participants. We omitted 34 individuals who did not meet inclusion criteria. Approximately 15% had at least one missing value. Ipsative mean imputation was used to estimate missing values; however, cases in which mean scale values for the PSD or NSD could not be obtained given the amount of missing data (i.e., only one item was endorsed) were deleted ($n = 10$). We included the Inconsistency and Infrequency scales of the Attentive Responding Scale (ARS; Rogge & Maniaci, 2010) to screen for inattentive or random responding, and excluded 27 individuals because their scores indicated extreme inconsistency in responding or extreme endorsement of atypical responses. For the longitudinal analyses, we only retained cases that had at least three waves of completed CSI data ($n = 510$).[1]

The final baseline sample ($N = 1,656$) was predominantly female ($n = 1,564$). The mean age in years for participants and their partners was 27.5 ($SD = 6.5$) and 29.3 ($SD = 7.2$), respectively. The modal educational level was a college degree for both participants ($n = 708$) and their partners ($n = 624$). Many individuals did not report their

1. About half of the initial sample was interested in participating in the follow-up assessments. These respondents provided their e-mail address and were sent invitations to complete brief follow-up assessments at 3, 6, 9, 12, 15, and 18 months. Those participating in the follow-ups tended to be slightly older and tended to have slightly higher levels of education.

monthly salary ($n = 168$) or their partner's monthly salary ($n = 301$). Of the remaining participants, the median monthly income reported was $2,500 for themselves and $3,000 for their partners. The majority of participants were Caucasian (86.8%), with the remainder comprising 2.4% African American, 3.3% Asian American/Pacific Islander, 3.6% Latino/a, .8% Native American, .6% Middle Eastern, and 2.2% identifying themselves as Other. Five individuals did not report on their ethnicity; 261 individuals reported that their ethnicity differed from that of their partner. Average relationship length in months was 28.9 ($SD = 26.5$) for individuals in serious dating relationships ($n = 357$), 43.7 ($SD = 27.0$) for those engaged to be married ($n = 623$), and 78.4 ($SD = 67.0$) for married participants ($n = 676$). The majority of married participants were cohabiting (98%), as were those engaged to be married (72%). Only 36% of the individuals in serious dating relationships were living with their partner. For participants living with their partner, the average length of cohabitation in months was 40.3 (SD = 53.3). Twenty-seven individuals reported that they were currently in therapy for relationship problems.

Measures

Positive and negative relationship satisfaction. The PN-SMD was formed using seven highly precise and informative semantic differential items, as well as Kaplan's (1972) method for constructing the positive and negative semantic dimensions. Note that Kaplan's method for reconfiguring semantic differential items used a 4-point scale, whereas we used a 7-point response scale. Our approach is more consistent with the longer response scale on the PANQIMS, which we presently adopted to increase response variability on our measure (thereby enhancing the variance of the final scale). Items within each subscale were administered together, with each group of items prefaced with the following statement: "Considering only the **positive/negative qualities** *of your relationship and* **ignoring the negative/positive ones**, evaluate your relationship on the following qualities. My relationship is …" This wording is consistent with the approach used by Kaplan and for the PANQIMS. It helps ensure that participants evaluate the positive and negative aspects of their relationship separately rather than applying the positive and negative semantic labels to the relationship as a whole. Also, this may help reduce the pull toward response consistency across ostensibly opposite semantic descriptors. For similar reasons, the PSD and NSD subscales were separated by additional questionnaires that were presented at random. The PSD and NSD subscales ranged from 0 to 49, with higher scores representing more positive and negative evaluations of the relationship, respectively. The sample mean for the PSD was 41.8 ($SD =$ 7.4) and 6.3 ($SD = 9.1$) for the NSD. The α coefficient was .95 for both the PSD and NSD. The PN–SMD items are presented in the Appendix.

One-dimensionalrelationship satisfaction. We used the 16-item version of the CSI. Using IRT analysis, Funk and Rogge (2007) identified these items as providing the most information about relationship satisfaction from a pool of 176 items sampled mostly from 12 predominant measures of relationship satisfaction (e.g., MAT). Scores ranged from 3 to 81; higher scores represent greater satisfaction. The mean at the initial assessment was 66.9 ($SD = 13.3$); alpha coefficients across time points were all in the acceptable range.

Negative and positive affectivity. The Eysenk Personality Questionnaire–Neuroticism (EPQ–N; Eysenk & Eysenk, 1975) comprises 23 dichotomously scored items that measure negative affectivity (e.g, "Do you ever feel 'just miserable' for no reason?"). Scale scores ranged from 0 to 23, had a mean value of 10.5 ($SD = 5.3$) and $\alpha = .85$; higher scores indicate greater negative affectivity. We included 12 items from the Anhedonic Depression (AD) subscale of the Mood and Anxiety Symptom Questionnaire (MASQ; Watson & Clark, 1991) as our measure of positive affectivity. Items on the MASQ–AD, which are rated on 5-point Likert scale ranging from 0 (*not at all*) to 4 (*extremely*), reference positive affect and cognition (e.g., proud of myself, felt

optimistic, felt cheerful). The sample mean was 29.6 ($SD = 8.6$); $\alpha = .92$.

Perceived social support. We assessed this construct with the Kessler Perceived Support Scale (KPSS; Kessler, Kendler, Health, Neale, & Eaves, 1992). Respondents were asked to rate their partner's support on a 4-point scale ranging from 1 (*not at all*) to 4 (*a great deal*) across three main stems: listen to you if you need to talk about your worries or problems; understand the way you feel and think about things; go out of their way to help you if you really need it. We added two additional questions to increase the variance of the scale (accept you totally, including both your worst and your best points; help you feel better when you are feeling stressed-out or down-in-the-dumps). Scores ranged from 0 to 15 with a sample mean of 13.3 ($SD = 2.7$); $\alpha = .89$.

Hostile conflict. Hostile and attacking conflict behavior was assessed using the Aversive Interaction Scale (AIS; Rodriguez & Rogge, 2012). Participants reported the frequency of their partners' behavior over the last year on 16 items (e.g., used profanity, shouted or yelled, ignored partner, etc.). Responses to these items were summed so that higher scores indicated higher levels of partner hostile conflict behavior. Scale scores ranged from 0 to 320, the average AIS score was 40.8 ($SD = 58.3$); $\alpha = .95$.

Sexual satisfaction. We used a seven-item version of the Index of Sexual Satisfaction (ISS; Hudson, Harrison, & Crosscup, 1981), which measures sexual satisfaction using evaluative statements (e.g., "My sex life is very exciting") rated on a 5-point Likert scale ranging from 0 (*rarely or none of the time*) to 4 (*most or all of the time*). The mean ISS score was 18.7 ($SD = 5.3$); $\alpha = .79$.

Results

Preliminary analyses

Univariate and multivariate outliers were retained in the set if their omission did not substantially alter the pattern of findings; otherwise, they were corrected with nonlinear transformations of the offending variables using the guidelines presented by Cohen, Cohen, West, and Aiken (2003). All variables were normally distributed and linearly related before the analysis. Inspection of residuals indicated potential heteroskedasticity in the regression models. To correct for this, we employed the heteroskedasticity-consistent standard error estimate procedure developed by Cribari-Neto (2004), and computed the analyses using the SPSS macro and syntax provided by Hayes and Cai (2007). Similarly, Levene's test indicated significant violation of the homogeneity of variance assumption across the analysis of variance (ANOVA) models. For these analyses, we used the Welch and Games–Howell tests for the omnibus analyses and planned comparisons, respectively, as these procedures do not require this assumption (see Games & Howell, 1976; Meyers & Well, 1995).

The factorial structure of the PN-SMD

Model comparisons. We conducted multigroup CFAs with relationship status as the grouping variable. We evaluated model fit using the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). CFI values at or higher than .90, RMSEA lower than .07, and SRMR values of less than .08 represent good fitting models (Tabachnick & Fidell, 2006). As χ^2 is sensitive to sample size, the χ^2/df criterion was used as a better indicator of fit when comparing across models. We also compared model fit using Akaike's information criteria (AIC; Akaike, 1987), with lower AIC values representing better fitting models. Fit statistics for each model are displayed in Table 1.

Our hypothesis that an oblique two-factor model would provide the best fit for the observed data was confirmed. Specifically, the two-factor oblique model was a reasonable fit to the data according to the CFI, RMSEA, and SRMR, whereas neither the one-factor nor the orthogonal two-factor model met any of the criteria for satisfactory fit. The performance of these latter two models was especially poor

Table 1. *Indices of fit for the multigroup confirmatory factor analyses using the PN–SMD*

	Indices of model fit				
Model	χ^2/df	CFI	RMSEA	SRMR	AIC
One factor	87.03	.08	.23	.66	21,801.7
Two-factor orthogonal	11.95	.89	.08	.39	2,928.1
Two-factor oblique	8.66	.93	.068	.03	2,198.1
Bifactor	11.82	.89	.08	.39	2,908.2

Note. $N = 1{,}656$. PN–SMD = positive and negative semantic differential; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; AIC = Akaike's information criteria. Adequate fit is indicated by CFI values > .90 and RMSEA and SRMR values below .07 and .08, respectively. Smaller χ^2/df and AIC values indicate better fitting models.

with respect to the SRMR fit index. Similarly, according to AIC values, the two-factor oblique model was a comparatively better fit to the data than was either the orthogonal two-factor or one-factor models. Taken together, these findings indicate that items measuring attitudes toward the positive and negative aspects of the relationship pertain to separate but related constructs.

However, given the high correlation between the latent PSD and NSD dimensions, we also tested the fit of a bifactor model (see Chen, West, & Sousa, 2006). This model included a third latent variable representing the common variance between the positive and negative dimensions (i.e., global satisfaction). In particular, we specified causal paths between this general factor and the items on both the PSD and NSD (items on each subscale still loaded separately onto their respective latent constructs). The fit for the bifactor model was not adequate, however, and comparatively worse than the two-factor oblique structure according to AIC values (see Table 1). As a set, the CFA findings suggest that individuals indeed hold distinct positive and negative feelings toward their romantic relationships. Although those attitudes are related to each other, the results suggested that they can be measured separately, and their correlation occurs for reasons external to the measurement model. The parameter estimates for the oblique two-factor model are presented in Figure 1.

Factorial invariance. On the basis of Byrne's (2010) recommendations, we examined if the results for the oblique two-factor model were invariant across relationship type (i.e., married, engaged, or dating). We found that all the constrained models provided a good fit and the estimates for the CFI, RMSEA, and SRMR did not appreciably differ between them (ranges = .90 to .93 for the CFI, .06 to .07 for the RMSEA, and .03 to .08 for the SRMR), suggesting that a single model may account for the associations among observed and latent variables across samples, and that the statistical differences across samples do not represent meaningful or systematic variation. We found a similar pattern when testing for factorial invariance across gender; the two-factor model was not invariant when constraining factor loadings and regression weights but nonetheless provided a good overall fit to the data. These results suggest that both men and women at various stages in their romantic relationships demonstrate the same tendency to hold separate (yet correlated) positive and negative attitudes about their relationships. As a result, the PN-SMD scale can be expected to operate similarly across a wide range of individuals in relationships.

Convergent and criterion-related validity of PN–SMD

Turning to the bivariate correlations (Table 2); the PSD and NSD not only demonstrated associations in the expected directions but also demonstrated some specificity from one another. Specifically, the PSD demonstrated

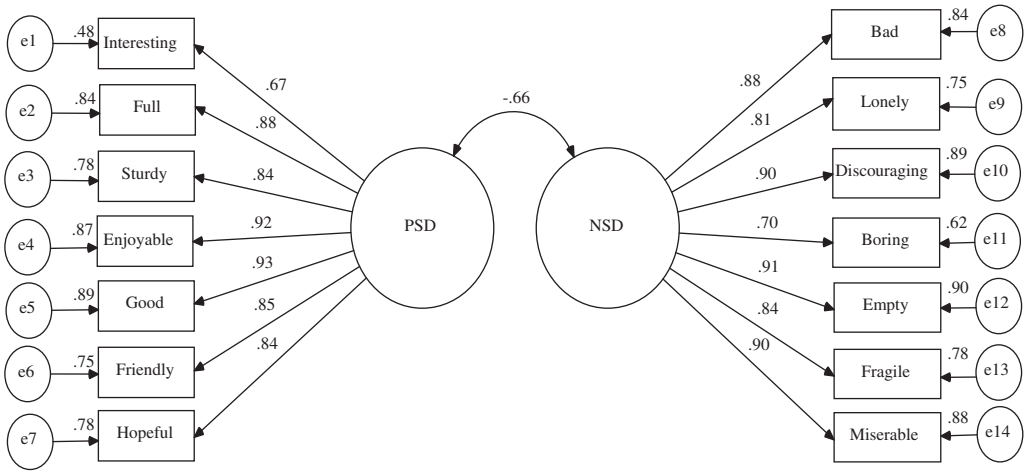

Figure 1. Parameter estimates for the oblique two-factor model of the PN−SMD. PN-SMD = positive and negative semantic differential; PSD = positive semantic dimension; NSD = negative semantic dimension. The items in the model appear in the sequence in which they were administered, beginning and ending with "Interesting" and "Hopeful," respectively, for the PSD, and "Bad" and "Miserable," respectively, for the NSD.

correlations that were greater in absolute magnitude to positive factors (social support and sexual satisfaction), whereas the NSD demonstrated correlations that were greater in absolute magnitude to negative factors (hostile conflict and negative affectivity). We tested whether the correlations between the PSD and NSD scores and the criteria were significantly different (in terms of absolute magnitude) using Meng, Rosenthal, and Rubin's (1992) method for comparing correlated correlations (see last row of Table 2). Our results indicate that the PSD correlated with social support and sexual satisfaction to a greater extent than did the NSD, whereas the NSD yielded stronger correlations with hostile conflict and neuroticism. These results therefore suggest that positive and negative feelings toward a

Table 2. *Convergent and criterion-related validity of the positive and negative semantic differential using partial correlations*

Correlations	Predictors			Criterion			
Variables	1	2	3	KPSS	AIS	ISS	EPQ-N
PSD	—	−.62	.71	.71	−.41	.46	−.25
NSD		—	−.58	−.64	.54	−.38	.33
CSI			—	.57	−.46	.43	−.27
Differences in correlation magnitude between PSD and NSD across criterion measures							
PSD−NSD$_{(z\text{difference})}$			8.51	4.60	−6.67	3.64	−3.89

Note. N = 1,656. PSD = positive semantic dimension; NSD = negative semantic dimension; CSI = Couples Satisfaction Index; KPSS = Kessler's Perceived Support Scale; AIS = Aversive Interaction Scale; ISS = Index of Sexual Satisfaction; EPQ−N = Eysenk Personality Questionnaire−Neuroticism. The coefficients represent partial correlations; gender, age, partner's age, relationship length, cohabitation status, and relationship type were statistically controlled. The differences in absolute magnitude of the correlations between the PSD and NSD and the criterion variables were evaluated using z tests, in accordance with Meng and colleagues' (1992) approach. All correlations and z tests were significant at $p < .01$ (two-tailed).

relationship might not only be empirically distinct (as suggested by the CFA) but might also have meaningfully different correlates and potentially link to different underlying processes or antecedents.

Incremental validity of PN–SMD

Model specification. We used hierarchical multiple regression to evaluate the incremental validity of the PN–SMD. We entered the following control variables into the first step: gender, age, partner's age, relationship length, cohabitation status, and relationship type. We entered the PSD and NSD in the second step to evaluate the incremental validity of each semantic dimension over the other, as well as to evaluate the overall explanatory power of both in tandem—beyond the control variables—according to change in R^2. We then entered in our one-dimensional relationship satisfaction measure (i.e., the CSI) in Step 3, followed by positive and negative affectivity (i.e., the MASQ–AD and EPQ–N, respectively) in Step 4. Semipartial correlation coefficients (sr) for the PN–SMD subscales in Steps 3 and 4 indicate, respectively, their incremental validity over a one-dimensional measure of satisfaction alone or in tandem with positive and negative affectivity. In accordance with Hunsley and Meyers (2003), we set the minimum criteria for meaningful incremental validity to $sr = .15$.

One-dimensional relationship satisfaction. The results—displayed in Table 3—support our hypotheses. Both the PSD and NSD demonstrated significant and distinct predictive associations with the four constructs examined, suggesting that both dimensions provide unique insight into these processes. As hypothesized, the PSD demonstrated significantly stronger and unique associations than the NSD when predicting sexual satisfaction and perceived partner supportiveness (demonstrating incremental validity). Likewise, the NSD demonstrated stronger and unique associations when reports of conflict behavior and negative affectivity were the criterion. The CSI only accounted for an additional 1% of variance in each of these constructs when added in Step 3, failing to yield incremental validity over the PN–SMD in any of the models (i.e., srs $<.15$). Moreover, the PSD and NSD retained their incremental validity (i.e., srs $>.15$) when both one-dimensional satisfaction (Step 3) and positive and negative affectivity were entered into the model (Step 4).

Considered together, our findings indicate (a) an empirical distinction between evaluations of the positive and negative aspects of the relationship (corroborating our bivariate findings), (b) that separate measures of these evaluative judgments confer information pertinent to relationship functioning otherwise missed by a one-dimensional scale, (c) even a precise one-dimensional measure like the CSI afforded no additional explanatory gains, and (d) the unique predictive validity of positive and negative relationship satisfaction is not reducible to potential overlap with positive or negative affectivity.[2]

Distinguishability between ambivalent and indifferent participants: Cross-sectional analyses

Using the PN–SMD, we attempted to replicate Fincham and Linfield's (1997) finding that mean differences in the criteria would exist between participants categorized

2. Two subsamples (ns = 835 and 821) were randomly generated, and the incremental validity analyses were rerun separately for each subsample in order to cross-validate the findings from the full sample. Using the CSI as the comparison, the sr coefficients were highly stable for hostile conflict and sexual satisfaction, with values differing from the full sample estimates by no more than approximately .01 in either direction. Social support evidenced a wider range of sr values for the PSD (.27 to .33) and NSD (−.18 to −.26), but the pattern of findings was consistent with those from Table 2. With regard to negative affectivity, the sr values for the PSD differed from the full sample estimates by no more than .01 in either direction, although the NSD demonstrated a wider range (−.15 to −.23), all effects were above Hunsely and Meyers's (2003) criterion for incremental validity. Similar findings emerged when entering positive and negative affect in Step 4. However, in one and the other subgroup, the sr value for the PSD and NSD in predicting sexual satisfaction ($sr = .14$) and negative affectivity ($sr = .11$), respectively, dropped below our criteria (though remained statistically significant predictors).

Table 3. *Incremental validity of the PSD and NSD subscales beyond a one-dimensional measure of relationship satisfaction and positive and negative affect*

	Social support		Hostile conflict		Sexual satisfaction		Negative affect	
	ΔR^2	*sr*	ΔR^2	*sr*	ΔR^2	*sr*	ΔR^2	*Sr*
Step 2	.54**		.29**		.23**		.11**	
PSD		**.39****		−.07**		**.28****		−.06*
NSD		**−.24****		**.37****		−.13**		**.22****
Step 3	.001*		.01**		.01*		.01**	
PSD		**.31****		−.01		**.17****		−.01
NSD		**−.22****		**.34****		−.09**		**.20****
CSI		.03*		−.08**		.12**		−.08**
Step 4	.001		.01**		.01*		.11**	
PSD		**.30****		−.03		**.16****		.05*****
NSD		**−.22****		**.32****		−.08**		**.16****
CSI		.03**		−.08**		.11**		−.04*****
Positive affect		.02		.07**		.06**		**−.34****
Negative affect		.03*		.10**		.03		**—**

Note. $N = 1{,}656$. PSD = positive semantic dimension; NSD = negative semantic dimension; CSI = Couples Satisfaction Index. Control variables were entered in the first step but were not tabulated in the interest of space. The criterion for incremental validity was a semipartial correlation (*sr*) of .15 or greater; estimates that met this criterion are highlighted in bold. Significance values were corrected for heteroskedasticity across models.
$^{*}p < .05$. $^{**}p < .01$ (two-tailed).

(using median splits) as satisfied, dissatisfied, ambivalent, or indifferent,[3] and that differences would exist between the latter two groups for all variables except the CSI. Because we tested five separate dependent variables, we set the p value per contrast at .01 to control for error inflation. We found that (a) the omnibus test revealed significant group differences in mean values across all of the criteria; values ranged from $F_{\mathrm{w}}(3, 654.0) = 67.5$ to $F_{\mathrm{w}}(3, 596.8)$ 400.6, p's <.01; (b) ambivalent and indifferent participants were indistinguishable based on their CSI scores; $\upsilon_{i,j}$[4] $= 1.89$ ($SE = .63$), *ns*; and, relative to indifferent participants, (c) ambivalent individuals reported greater negative affectivity, $\upsilon_{i,i} = 2.52$ ($SE = .46$) and partner's conflict behaviors, $\upsilon_{i,j} = 20.42$ ($SE = 4.57$), ps $< .01$. As shown graphically in Figure 2, the significant planned contrasts indicate that ambivalent individuals reported greater negative affectivity and more conflict behaviors than indifferent individuals, despite the fact that these two groups endorsed similar CSI scores on average. When analogous interaction terms between PSD and NSD were added to the regressions presented in Table 3 (extending these median split analyses to their continuous forms), the same pattern of findings emerged across the criterion variables.

Taken together, the PN–SMD identified two classes of individuals (those holding ambivalent or indifferent attitudes toward their relationships) that were substantively different from one another (as well as from distressed and satisfied individuals) on specific relationship processes but were indistinguishable at the cross-sectional level on the basis of a one-dimensional satisfaction index. Furthermore, our results demonstrate

3. We separated participants into groups as follows: individuals scoring above the median on the PSD and at or below the median on the NSD scales were categorized as satisfied, those scoring at or below the median on the PSD and above the median on the NSD were classified as dissatisfied, those scoring above median on both subscales were classified as ambivalent, and those scoring at or below the median on both subscales were classified as indifferent.
4. This symbol signifies the difference between the ith and jth means, which in this case are the means for the ambivalent and indifferent groups, respectively.

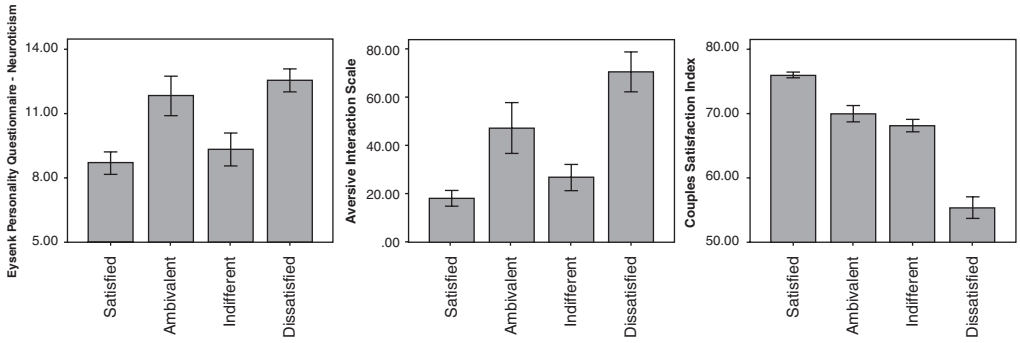

Figure 2. Mean differences between satisfied, dissatisfied, ambivalent, and indifferent individuals with respect to negative affectivity (Eysenk Personality Scale–Neuroticism), conflict behaviors (Aversive Interaction Scale), and relationship satisfaction (Couples Satisfaction Index) using the Games–Howell procedure. A p value of .01 was used for each planned comparison to control for error inflation.

that the same overall pattern of findings observed when using the PANQIMS can also be detected when using the PN–SMD.

Distinguishability between ambivalent and indifferent participants: Longitudinal analysis

We examined whether distinct trajectories of global relationship quality over time would emerge across the four main groups (i.e., satisfied, dissatisfied, ambivalent, and indifferent). Of particular interest was whether individuals identified as ambivalent (high PSD and high NSD) at the first assessment would display steeper declines in CSI scores over time relative to those identified as indifferent (low PSD and low NSD). We tested our hypotheses using a series of slope-intercept hierarchical linear models (HLM) where individual differences were modeled at Level 2 and repeated assessments across time were modeled at Level 1. Both initial status (intercept) and linear change over time (slope) were set as random effects at Level 2 to allow individuals to vary in their linear trajectories of one-dimensional global relationship satisfaction. Instead of forming groups from median splits in these analyses, we included continuous interaction terms between PSD and NSD. Following standard practice when graphing significant interactions (see Curran, Bauer, & Willoughby, 2006), we then used 1 *SD* above or below the mean on the PSD and NSD to represent the various groups (e.g., +1 *SD* on both PSD and NSD would represent ambivalent individuals, whereas −1 *SD* on both PSD and NSD would represent indifferent individuals). Prior to analysis, we standardized the criterion and the predictors (except for time) to facilitate interpretation of the parameter estimates (see Table 4).

Model A. A slope-intercept model without any predictors (unconditional growth model) suggested that global relationship satisfaction scores tended to drop over the 18 months of the study. The pseudo R^2 statistic (see Snidjers & Bosker, 1994) suggested that this linear trend accounted for roughly 13% of the variance in scores over time (when compared to a fully unconditional model).

Model B. We then entered the initial reports of positive (PSD) and negative (NSD) relationship attitudes (along with a PSDxNSD interaction term) into the portion of the model predicting global relationship satisfaction at the start of the study. The results suggested that higher baseline positive attitudes toward the relationship corresponded to higher levels of initial relationship satisfaction, whereas higher baseline negative attitudes toward the relationship associated with lower levels of initial relationship satisfaction. Initial positive

Table 4. *Results for the multilevel models examining the effects of initial PSD and NSD status on CSI intercepts and slopes over time*

Fixed effects	Parameter		Model A	Model B	Model C	Model D
Initial status	Intercept	γ_{00}	.09 (.05)	.02 (.03)	.02 (.03)	.04 (.03)
	PSD	γ_{01}		.57 (.04)*	.59 (.04)*	.62 (.03)*
	NSD	γ_{02}		−.28 (.03)*	−.29 (.03)*	−.33 (.03)*
	PSD*NSD	γ_{03}		−.03 (.02)	.03 (.02)	.02 (.02)
Linear change over time	Intercept (time)	γ_{10}	−.21 (.03)*	−.22 (.03)*	−.18 (.04)*	−.13 (.03)*
	PSD	γ_{11}			−.10 (.04)*	.01 (.04)
	NSD	γ_{12}			.03 (.04)	−.07 (.03)*
	PSD*NSD	γ_{13}			.02 (.03)	.02 (.02)
Shifts within each wave	PSD	γ_{20}				.38 (.02)*
	NSD	γ_{21}				−.33 (.02)*
Random effects						
Level 1	Within-person	σ^2_{ϵ}	.23 (.01)*	.23 (.01)*	.23 (.01)*	.16 (.01)*
Level 2	In initial status	σ^2_0	.74 (.06)*	.12 (.02)*	.12 (.02)*	.09 (.01)*
	In rate of change	σ^2_1	.15 (.03)*	.16 (.02)*	.15 (.03)*	.08 (.02)*
	Covariance	σ^2_{01}	−.02 (.03)	.06 (.02)*	.06 (.02)*	−.01 (.01)

Note. $n = 510$. PSD = positive semantic dimension; NSD = negative semantic dimension; CSI = Couples Satisfaction Index. All the modeled variables except for time were standardized before analysis. The fixed and random effects were evaluated for statistical significance using t tests and Wald Z values, respectively.
*$p < .05$ (two-tailed).

and negative attitudes toward romantic relationships accounted for 84% of the variance in global relationship satisfaction at the start of the study. Although the interaction term was not significant, we nonetheless plotted the intercepts for each group (Figure 3a) to show that findings from the present analysis mirror the median-split results presented before.

Model C. We then added PSD, NSD and an interaction between the two as predictors for linear change in global relationship satisfaction over time. The results suggested that higher initial positive attitudes toward the relationship were associated with slight declines in relationship satisfaction over time, accounting for 6% of the variance in individual CSI trajectories. This may suggest that changes in one-dimensional satisfaction for ambivalent (and satisfied) individuals resulted because their initially more favorable evaluations of the relationship's positive characteristics were diminishing over the 18-month course of the study. Neither NSD nor the interaction term demonstrated significant associations with linear change in relationship satisfaction. Taken together, these findings suggest that (a) the more critical determiner of the slope of one-dimensional satisfaction was high versus low baseline PSD scores and (b) this effect did not differ across those with high versus low negative relationship evaluations. In other words, individuals with high initial PSD scores (i.e., satisfied and ambivalent) showed steeper declines irrespective of their evaluative judgments toward the negative aspects of their relationship. To illustrate, the regression weights were used to generate prototypic change trajectories for dissatisfied, indifferent, ambivalent, and satisfied participants (see Singer & Willet, 2003), which are displayed in Figure 3b.

Model D (exploratory analyses). To further explore unique links between positive–negative relationship attitudes and global relationship satisfaction across time, we built a final model in which we entered shifts in positive and negative relationship attitudes at each wave of assessment (above and below one's own initial levels of each attitude)

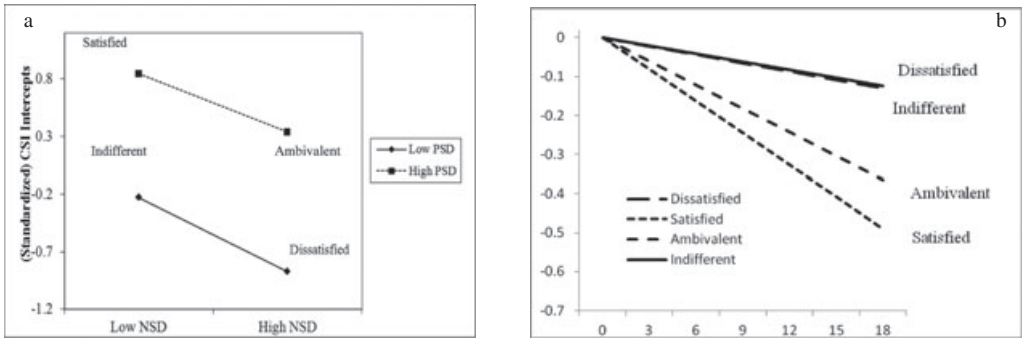

Figure 3. Prototypic Couples Satisfaction Index (CSI) intercepts (a) and trajectories (b) for dissatisfied, indifferent, ambivalent, and satisfied participants over 18 months. CSI scores were square root transformed before the analysis.

as time-varying covariates to predict corresponding shifts in global relationship satisfaction in those same waves of assessment. The results suggested that spikes in positive relationship attitudes (above one's own initial levels) tended to be associated with corresponding spikes in global relationship satisfaction in those same waves. Similarly, spikes in negative relationship attitudes tended to be associated with corresponding dips in global relationship satisfaction. Together, shifts in positive and negative relationship attitudes accounted for an additional 46% of the variance in global relationship satisfaction across time. This suggests that shifts in global relationship satisfaction can result from fluctuations in either positive or negative attitudes toward the relationship. As the underlying processes influencing shifts in positive and negative attitudes are likely to be distinct, these results once again highlight how the use of a one-dimensional measure of global relationship satisfaction might serve to obscure meaningful distinctions in attitude change over time (e.g., the difference between drops in negative attitudes and increases in positive attitudes would be obscured). Also notable, initial PSD values no longer significantly predicted CSI slopes when time-varying PSD scores were added to the model. This implies that subsequent changes in positive relationship attitudes over time mediate the association between their baseline values and CSI slopes. That is, higher baseline PSD scores are perhaps more likely to depreciate over time, which, in turn, underscore contemporaneous declines in one-dimensional satisfaction.

Discussion

The results of Study 1 provide overall support for the two-dimensional structure of relationship satisfaction. With respect to our CFA findings, the poor fit for the one-dimensional model indicated that conceptualizing relationship satisfaction as a single latent construct failed to capture its underlying dimensionality. Notably, however, the correlation between the positive and negative dimensions was fairly high ($r = -.66$), suggesting that the two dimensions share roughly 44% of their variance. Although this may be suggestive of an overarching evaluative judgment (a single common factor) with separable positive and negative attitude domains nested therein (i.e., a bifactor model; see Chen et al., 2006), the present data did not support this contention. Rather, the data supported a two-factor solution in which evaluations of the positive and negative relationship characteristics correlate for some reason outside of the measurement model.

One possibility is that the association observed between the PSD and NSD reflects the correlation between their respective evaluative objects. Relationships higher in positive characteristics (e.g., social support) may tend to be lower in negative characteristics (e.g., conflict), and so the evaluations

of these separate domains will be correlated but nonetheless distinct attitudes. Also, the association between the PSD and NSD is similar to that observed when assessing positive and negative aspects of other constructs. For example, common two-dimensional scales assessing positive and negative mood states (e.g., the Positive and Negative Affect Scale [PANAS]; Watson, Clark, & Tellegen, 1988) demonstrate comparable associations between the two scales. In any case, the median split and longitudinal HLM analyses suggest that—although the positive and negative dimensions share variance (as might be expected because they reflect attitudes toward the same relationship)—they provide distinct insights into relationship processes. Moreover, detecting this nuance may have important clinical implications and theoretical relevance to developmental models of relationship distress.

The regression analyses also supported a two-dimensional structure for relationship satisfaction. If positive and negative attitudes were truly bipolar, then responses on one dimension (e.g., "My relationship is 'completely' boring") would be the exact inverse of responses on the other (i.e., "My relationship is 'not at all' interesting"), and variance in the criteria explained by either dimension would completely overlap. This was not the case, however; either positive or negative relationship evaluations (or both) uniquely associated with the criteria when controlling for overlapping variability between them. Furthermore, they accounted for variance in our criteria beyond a one-dimensional measure of relationship satisfaction. This latter finding is consistent with a two-dimensional formulation. Specifically, forcing individuals to evaluate the positive and negative aspects of their relationship using one global dimension (ranging from positive to negative) will result in the loss of important information about relationship functioning. This appears to be particularly relevant for discriminating between indifference and ambivalence toward the relationship. Specifically, we found that ambivalent and indifferent individuals reported differences in negative affectivity and conflict behaviors but were indistinguishable in terms of CSI scores at the baseline assessment. This implies that a one-dimensional approach may treat qualitatively different underlying attitudes as equivalent, thereby obscuring the aspects of relationship functioning that underlie these distinct evaluative judgments.

We also found that initially endorsing more favorable evaluations of a relationship's positive characteristics predicted steeper declines in one-dimensional relationship satisfaction. However, the impact of higher initial positive evaluations was potentially mediated by subsequent changes in these same attitudes. That is, after controlling for change in positive relationship attitudes across subsequent assessment waves; the effect of their baseline values on change in one-dimensional satisfaction was no longer significant. This implies that highly positive relationship evaluations were more likely to attenuate over time, which, in turn, emerged as contemporaneous decreases on a one-dimensional relationship satisfaction measure. However, we found that evaluations of the relationship's negative characteristics at each assessment wave were also longitudinally relevant. As such, reported declines in global relationship satisfaction may variously reflect changes in either or both positive and negative relationship evaluations. Such findings highlight the interpretive difficulties inherent to one-dimensional measures of relationship satisfaction. Specifically, the underlying nature and course of a relationship's evaluation is apparently more complex than is evident when using a one-dimensional scale. By comparison, the interplay of the positive and negative evaluative dimensions may provide the most comprehensive understanding of relationship functioning.

Study 2

Study 2 replicated the convergent, criterion-related, and incremental validity of the PN-SMD. We conducted Study 2 in a laboratory setting, as opposed to online, and we administered alternative measures of partner supportiveness, conflict, and negative affectivity. In

addition, we expanded the set of criterion variables to include maladaptive relationship attributions for partner behavior. The PN–SMD was again compared to the CSI for the purpose of replication. However, Study 2 also extended the findings of Study 1 in three ways. First, we examined the incremental validity of the PN–SMD over the original semantic differential. Comparing the PN–SMD with its one-dimensional counterpart more directly assesses the contribution of a two-dimensional assessment methodology of relationship satisfaction, as important psychometric differences between the measures are held constant (e.g., semantic content). Second, we explored the incremental validity of the PN–SMD over the PANQIMS, as use of the former is only justified if its validity exceeds that of its predecessor (Haynes & Lench, 2003; Hunsley & Meyers, 2003). Third, based on Kaplan's (1972) attitude component model, we created indices of ambivalence (i.e., the extent to which PSD and NSD are equal) and polarization (i.e., the magnitude of the difference between the PSD and NSD dimensions). We examined whether this method of modeling ambivalence toward the relationship would meaningfully improve the prediction of the criterion variables beyond the effects captured by more polarized (i.e., unilaterally favorable versus unfavorable) attitudes, which are effectively similar to a typical one-dimensional measure of relationship satisfaction.

Method

Participants

Undergraduate psychology majors at a university in the Southeastern United States participated in the study. A total of 89 participants were recruited through flyers placed on campus, announcements made in undergraduate classes, and through an online recruitment system used by the Department of Psychology. Participants were at least 19 years of age and in a monogamous dating relationship for at least 6 months (75.3%), married (4.5%), or engaged to be married (16.9%). The sample comprised 20 men and 69 women. The average age in years for participants and their partners was 21.3 ($SD = 1.9$) and 22.1 ($SD = 2.8$), respectively. The average relationship length was 24.9 months ($SD = 20.3$). The sample predominantly comprised Caucasians (78.7%); the racial demography of the remaining participants was as follows: 18.0% African American, 1.1% Asian/Pacific Islander, and 2.2% Other. By their report, 51.7% were currently working. The median monthly income for participants and their partners was $400 and $625, respectively. Ten participants were cohabiting with their partner, whereas 17 were in a long-distance relationship. Two individuals were in couples therapy.

Measures

Consistent with Study 1, we administered the same measures of social support (KPSS; $M = 23.9$, $SD = 5.5$; $\alpha = .87$), negative affect (EPQ-N; $M = 31.9$, $SD = 4.7$; $\alpha = .82$), positive affect ($M = 41.1$, $SD = 7.1$; $\alpha = .85$), and the PN-SMD's positive (PSD; $M = 40.4$, $SD = 7.4$; $\alpha = .94$) and negative (NSD; $M = 5.7$, $SD = 6.7$; $\alpha = .88$) subscales.[5] However, for the purposes of systematic replication, we used alternative measures of several variables from Study 1 (presented below).[6]

Positive and negative relationship satisfaction. The PANQIMS items were reconfigured to assess relationship (as opposed to marital) satisfaction; all references to marriage or spouse were replaced with relationship or partner, respectively (also see Mattson et al., 2007). Item responses are summed within each dimension; higher scores on the PMQ and NMQ reflect greater positive and negative relationship satisfaction, respectively. Both the PMQ and NMQ subscales were internally consistent, yielding reliability coefficients of .79 and .89, respectively. The mean PMQ and NMQ scores, respectively, were 18.3 ($SD =$

5. For the KPSS we used a 6-point Likert-type scale ranging from 1 (*not at all*) to 6 (*completely*), and added a question regarding overall satisfaction with partner supportiveness.
6. Although we administered a one-item measure of sexual dissatisfaction, these data were not analyzed because of insufficient variability across participants.

2.6) and 6.6 ($SD = 4.4$). PMQ scores ranged from 9 to 21, whereas NSD scores ranged from 0 to 21.

One-dimensional relationship satisfaction. We used the 16-item version of the CSI and the 7-item semantic differential (see Funk & Rogge, 2007). Scores on the CSI ranged from 44 to 97, with a mean score of 80.9 ($SD = 13.2$). Semantic differential scores ranged from 15 to 42 and had a mean of 36.0 ($SD = 5.7$). The CSI and semantic differential scales yielded αs of .96 and .90, respectively. Note that the CSI contains 6 of the 7 semantic differential items, but the 7-item semantic differential was examined separately to more specifically compare the one- versus two-dimensional methods.

Social support. As an alternative measure of perceived partner support, we administered the 12-item Perceived Responses to Capitalization Attempts scale (PRCA; Gable, Reis, Impett, & Asher, 2004). This scale measures a partner's responses to the sharing of positive events in the respondent's life (e.g., "My partner reminds me that most good things have their bad aspects as well"). Participants rate each statement using a 7-point Likert-type scale ranging from 0 (*not at all true*) to 6 (*very true*). We used Gable, Gonzaga, and Strachman's (2006) method for calculating a total score index for which higher scores indicate higher perceived levels of positive versus negative responses to capitalization attempts. Scores on the PRCA ranged from −9.7 to 6.0, and yielded a mean of .95 ($SD = 3.6$) and an α coefficient of .84.

Conflict behaviors. We assessed conflict behaviors using the Marital Coping Inventory–Conflict scale (MCI–C; Bowman, 1990). This measure requires participants to identify the most serious recurring problem in their relationship, followed by 15 items that assess hostile conflict behaviors (e.g., criticism) engaged in by the respondent and directed toward his or her partner with respect to this problem (e.g., "When I am dealing with the problem. . .I yell or shout at my partner"). Response scales ranged from 1 (*never*) to 5 (*usually*), with summed scores currently ranging from 32 to 75. Higher scores represent more hostile conflict behaviors. The mean for the MCI–C was 58.6 ($SD = 10.3$); $\alpha = .92$.

Neuroticism. In addition to the EPQ–N, we used the General Distress (GD): Depressive Symptoms subscale of the MASQ, which contains 12 items that assess symptoms of depressed mood (e.g., "felt sad") and nonspecific symptoms of mood disorder (e.g., pessimism) using a 5-point Likert scale. Qualitative anchors ranged from *not at all* to *extremely*; higher scores on the omnibus index indicate higher levels of negative affectivity. Scores presently ranged from 12 to 52 and had a mean of 23.1 ($SD = 8.3$); $\alpha = .89$.

Relationship attributions. We used the Relationship Attribution Measure (RAM; Fincham & Bradbury, 1992) to assess maladaptive attributions for negative partner behaviors (e.g., "My partner is cool and distant on purpose rather than unintentionally"). In the interest of space, we used a truncated version of this scale containing 12 items that assess attributions across two domains: (a) your partner is cool and distant and (b) your partner is not paying attention to what you are saying. Items were summed to form an omnibus index, with higher scores indicating more maladaptive attributions for partner behavior. The mean RAM score was 29.4 ($SD = 9.6$); scores ranged from 11 to 52 and the coefficient α was .85.

Procedures

Participants were individually assessed in the laboratory. Research assistants provided informed consent forms to participants and reminded them that they could withdraw from the study at any time without penalty. Participants then completed the survey packet, which took approximately 25–30 min. Each participant was then offered a referrals list for couples counseling, debriefing forms, and contact information for the principal investigator. Participants were offered extra credit for their psychology courses in return for their participation.

Results

Preliminary analyses

Preliminary analyses showed univariate outliers, as well as significant skew and kurtosis, for negative affect (MASQ–GD only), negative evaluations of the relationship (both the NMQ and NSD), perceived responses to capitalization attempts, and the semantic differential. A square root transformation for these variables was required to achieve normality. Age, partner's age, and relationship length were also square root transformed. There was evidence for heteroskedasticity when the MASQ–GD was the criterion; as such, we used the heteroskedasticity-consistent standard error estimate procedure employed in Study 1 for this regression model (Table 4).

Convergent and criterion-related validity of the PN–SMD

We hypothesized that the findings of Study 1 would replicate using additional measures of relationship satisfaction and an expanded set of criteria (controlling for gender, age, partner's age, relationship length, cohabitation, and relationship type). This hypothesis was supported (Table 5). The PSD and NSD converged with both one-dimensional measures of relationship satisfaction (the CSI–16 and the semantic differential) and both the PANQIMS subscales in the expected directions, with medium to large effect sizes. Association with social support and negative affect (using the EPQ–N) were similar in magnitude to their corresponding estimates in Study 1. In addition, we found that positive and negative relationship attitudes significantly correlated with the added indices for negative affectivity (except for the PSD and MASQ–GD), hostile conflict, maladaptive attributions, and perceived responses to capitalization attempts (all in the anticipated directions). To facilitate the comparison between the PSD and NSD and their PANQIMS counterparts, we calculated z tests for differences between correlated correlations (Meng et al., 1992). These results—also displayed in Table 5—showed that the PSD correlated with social support and capitalization attempts to a significantly greater extent than did the PMQ, whereas the NSD was significantly more highly correlated with hostile conflict than was the NMQ. In addition to corroborating the results from the first study, these additional findings suggest that expanding the number and diversity of items used to assess the positive and negative semantic dimensions of relationship satisfaction provides for stronger correlations with relationship-relevant criteria.

Incremental validity of the PN-SMD

One- and two-dimensional relationship satisfaction measures. We hypothesized that PN–SMD subscales would demonstrate incremental validity over the CSI, the bipolar semantic differential, and the PANQIMS. We displayed our findings in Table 6. We found that the positive and negative subscales on the PN–SMD predicted unique and meaningful variability in the positive and negative aspects of the relationship, respectively, as compared to the CSI and bipolar semantic differential. This further suggests that researchers can obtain a deeper level of information about relationship satisfaction by separately assessing the positive and negative evaluative dimensions. The comparison between the bipolar semantic differential and PN–SMD makes this point most clearly; these measures contain the same semantic anchors but differ in their underlying dimensionality. The results also highlight the PN–SMD as significantly more valid than an existing measure of positive and negative attitudes toward the relationship. Although the NMQ provided incremental utility when predicting perceived responses to capitalization attempts and maladaptive attributions in Step 3, so did the NSD and, in the former case, the PSD. In other words, despite some support for the NMQ, the PN–SMD subscales were still the stronger and more consistent predictors overall. These findings were robust in Step 4, wherein we entered the MASQ–AD and MASQ–GD scales to measure positive and negative affect, respectively, and included the EPQ–N as an additional measure of the latter construct.

Table 5. *Convergent and criterion-related validity of the positive and negative semantic differential using partial correlations for Study 2 and z tests comparing correlations between the PN–SMD and PANQIMS subscales and the criteria*

	Predictors						Criterion					
	1	2	3	4	5	6	KPSS	PRCA	MCI–C	EPQ–N	MASQ–GD	RAM
PSD	—	−.52**	.86**	.81**	.54**	−.49**	.72**	.51**	−.37**	−.24*	−.05	−.38**
NSD		—	−.63**	−.61**	−.32**	.53**	−.56**	−.52**	.49**	.34**	.30**	.50**
CSI			—	.93**	.53**	−.52**	.74**	.53**	−.35**	−.29*	−.17	−.40**
SMD				—	.52**	−.44**	.66**	.51**	−.38**	−.34**	−.19	−.36**
PMQ					—	−.12	.26*	.18	−.26*	−.10	.05	−.22*
NMQ						—	−.48**	−.51**	.30**	.30*	.23*	.50**
Differences in correlation magnitude between PSD and NSD across criterion measures												
					PSD–PMQ		5.49**	3.47**	−1.13	−1.38	.97	−1.64
					NSD–NMQ		−.93	.11	2.03*	.40	.74	.61

Note. $N = 87$. PSD = positive semantic dimension; NSD = negative semantic dimension; CSI = Couples Satisfaction Index; SMD = semantic differential; PMQ = Positive Marital Quality; NMQ = negative marital quality; KPSS = Kessler's Perceived Support Scale; PRCA = Positive Responses to Capitalization Attempts; MCI–C = Marital Coping Inventory–Conflict; EPQ–N = Eysenk Personality Questionnaire–Neuroticism; MASQ–GD = Mood and Anxiety Questionnaire–General distress; RAM = Relationship Attribution Measure. Gender, age, partner's age, relationship length, cohabitation status, and relationship status were controlled. The differences in correlation magnitude between the PSD and NSD were evaluated using z tests, in accordance with Meng and colleagues' (1992) approach.

*$p < .05$. **$p < .01$ (two-tailed).

Table 6. *Steps 2 and 3 across hierarchical regression models examining the incremental validity (* sr*) of the PN–SMD over one- and two-dimensional measures of relationship satisfaction*

	Step 2			Step 3							
				Model 1				Model 2			
	R^2	PSD	NSD	R^2	PSD	NSD	CSI	R^2	PSD	NSD	SMD
Social support	.51**	**.47****	**−.21****	.02*	**.17***	−.13	.14*	.00	**.29****	**−.17***	.06
Responses to capitalization	.33**	**.28****	**−.29****	.00	.13	**−.24****	.05	.00	**.15***	**−.24****	.06
Hostile conflict	.21**	−.12	**.32****	.01	−.14	**−.32****	.09	.00	−.08	**.29****	.01
Negative affect (EPQ–N)	.10**	−.08	**.23***	.00	.00	**.18**	−.06	.00	.06	**.15**	**−.15**
Negative affect (MASQ–GD)	.10*	.12	**.31****	.01	**.16**	**.23***	−.12	.02	**.19**	**.23**	−.14
Maladaptive attributions	.23**	−.13	**.33****	.00	−.09	**.30****	.01	.00	−.11	**.32****	.04

	Step 3 Model 3			
R^2	PSD	NSD	PMQ	NMQ
.03	**.45****	**−.19****	**−.15***	−.02
.05	**.23****	**−.20***	−.09	−.17
.00	.08	**−.28****	.05	−.01
.01	−.05	**.17**	.03	.11
.02	.09	**.25***	.08	.11
.06*	−.02	**.21***	−.06	**.25****

Note. $N = 89$. PSD = positive semantic dimension; NSD = negative semantic dimension; CSI = Couples Satisfaction Index; SMD = semantic differential; PMQ = Positive Marital Quality; NMQ = Negative Marital Quality; EPQ–N = Eysenk Personality Questionnaire–Neuroticism; MASQ–GD = Mood and Anxiety Questionnaire–General distress. The criterion for incremental validity was a semipartial correlation (*sr*) of .15 or greater; estimates that met this criterion are highlighted in bold.

*$p < .05$. **$p < .01$ (two-tailed).

Exploratory analyses. As our samples disproportionately comprised women, we examined whether the incremental validity findings would be robust when only examining men. We combined the men from both samples ($n = 112$) and ran models using the EPQ–N (negative affectivity) and the KPSS (social support) as criteria (only these were measured consistently across studies). We added sample membership as an additional control variable. Our previous findings were generally robust: (a) the CSI did not predict unique variance in social support over the PN-SMD subscales; (b) evaluations of the relationship's negative characteristics significantly and uniquely predicted negative affectivity (sr = .20) when controlling for the CSI, $p < .01$; and (c) unique variance in social support was predicted by both negative ($sr = -.18$) and positive ($sr = .17$) relationship evaluations when controlling for positive and negative affect, ps <.01. In summary, although the magnitude of the association varied somewhat, the pattern of findings for our subsample of men was consistent overall.

Ambivalence as a predictor of individual and relationship functioning

We tested the incremental validity of a continuous index of ambivalence over one of polarization, which we computed using Kaplan's (1972) attitude component model. According to this model, the sum of the PSD and NSD dimensions can be thought of as the overall magnitude of evaluative sentiments—positive and negative—held toward the relationship (scores of [PSD + NSD] = 0 represent the absence of any evaluative sentiment about the relationship). The polarization and ambivalence indices represent the two subcomponents of the overall attitude. Polarization is the portion of the overall attitude that is unilaterally in either one direction (e.g., positive) or the other (i.e., negative), and is quantified by the absolute difference between the PSD and NSD scores (i.e., |PSD − NSD|). The ambivalence component represents the degree of opposing sentiments of equal magnitude and is computed by subtracting the polarization score from the magnitude of the overall sentiment (i.e., [PSD + NSD] − |PSD − NSD|). In other words, ambivalence occurs when some form of sentiment is held (i.e., not indifference) but is not distinctly positive or negative in direction.

We controlled for attitudinal polarization when examining ambivalence because any effects for the latter may actually represent those for polarization, as the two indices are likely negatively correlated. (The range of possible values for the polarization index necessarily decreases as ambivalence increases, and vice versa, when holding constant the total magnitude of the evaluative sentiment.) Using regression analysis, we found that greater levels of ambivalence were uniquely predictive of negative affect (using the MASQ–GD) and hostile conflict, srs = .26 and .19, respectively, ps <.05 (two-tailed). We also found that increases in ambivalence significantly predicted increases in relationship satisfaction, as measured by the CSI ($sr = .15$) and the bipolar semantic differential ($sr = .12$), ps <.05 (two-tailed), but only the former met Hunsley and Meyers's (2003) criteria for substantial incremental validity.

Discussion

The consistency in findings across studies is notable in light of the varying composition of the samples, methods of recruitment, and context of questionnaire administration. Findings that are robust across measures of the criteria and varying investigatory methods provide converging operations supporting their validity. We also compared the PN–SMD with its one-dimensional counterpart. As important psychometric properties across these measures were controlled, our findings suggest that assessing relationship satisfaction across two dimensions is incrementally more valid than doing so with a single bipolar continuum. Although the PN–SMD has two 7-point response scales (one for each dimension), whereas the bipolar semantic differential has only one, this difference cannot account for the increased validity of the two-dimensional model approach. Specifically, if semantic anchors truly belonged on a single bipolar dimension (i.e., were not independent), then responses to items on the NSD

(e.g., bad) and PSD (e.g., good) should be the exact inverse of each other and should correlate perfectly—in one direction or the other—with responses on the bipolar semantic differential (e.g., bad–good). As such, the PSD and NSD should not have explained any additional variance in the criteria over each other—let alone the bipolar semantic differential—unless they represented two independent attitudes.

The PN–SMD was also incrementally more valid than a different two-dimensional measure. However, the NMQ subscale did uniquely explain substantial variance in maladaptive attributions and partner's responses to capitalization attempts. It is possible that these findings occurred because the PANQIMS primarily samples attitudes about the partner, as opposed to the items on the PN–SMD that focus exclusively on the relationship. This may imply that attitudes toward the partner versus the relationship overall are partially separable and that both are associated with unique variability in the same criteria. Although this may be a fruitful area for subsequent research, this finding illustrates that the use of a measure combining potentially separable evaluative judgments may obscure the meaning of any associations with other variables.

Our results also showed that attitudinal ambivalence versus indifference about the relationship yields additional explanatory value. Specifically, when using Kaplan's (1972) method, attitudinal ambivalence accounted for unique variance in negative affectivity and hostile conflict when controlling for attitudinal polarization. These findings converge with the ANOVA results from Study 1, suggesting that negative affect and conflict are important correlates of relationship ambivalence. Perhaps, most interesting is that ambivalence also carried unique associations with one-dimensional measures of relationship satisfaction. This may suggest that individual differences in one-dimensional relationship satisfaction, in part, reflect differences in ambivalence rather than in satisfaction versus dissatisfaction per se. In any case, these findings highlight the increased utility of a two-dimensional approach, as deriving ambivalence scores requires separate indices for positive and negative relationship attitudes. Moreover, use of Kaplan's method may help open a link between relationship research and the broader literature on attitudinal ambivalence (e.g., Conner & Sparks, 2002).

General Discussion

According to Bradbury and colleagues (2000), the two-dimensional conceptualization of relationship satisfaction was one of most important advancements in the assessment of this construct (also see Fincham & Rogge, 2010). However, relatively few studies have used a two-dimensional assessment approach despite such endorsements and provided corroborating empirical support from subsequent validation studies (e.g., Mattson et al., 2007). The present findings provide additional evidence that measuring attitudes toward the relationship using two dimensions is incrementally more useful and valid, and that continued reliance on one-dimensional measures may hinder the detection or precise evaluation of important aspects of relationship functioning. Moreover, the presently developed measure derived from a clear theoretical domain of observables (i.e., SMD items) and an empirically guided approach to item selection and scale construction. Perhaps as a result, the PN–SMD emerged as more precise and easily interpretable than its one- and two-dimensional predecessors.

We also found evidence that attitudinal ambivalence versus indifference toward the relationship is an important distinction. Indeed, it is possible that such attitudes partially explain the modest association between relationship satisfaction and stability (i.e., whether or not the relationship ends). Specifically, attitudes represented at the midpoints of a one-dimensional scale (e.g., *somewhat unhappy*) may represent varying levels of ambivalence versus indifference, which may carry different associations with stability. This is consistent with our finding that ambivalent and indifferent individuals have different trajectories of one-dimensional satisfaction over

time. As such, they might even present differently in couples therapy and benefit from different methods of intervention. In any case, such interesting possibilities underscore the need for additional research examining ambivalent versus indifferent individuals, and highlight the PN–SMD's potential clinical utility over one-dimensional satisfaction measures that effectively lump these individuals into the same "moderately distressed" category.

Strengths and limitations

There were several notable strengths to our studies. First, we replicated our effects across two samples, controlling for several individual and dyadic variables, and using different operations of the criteria and alternative methods to capture attitudinal ambivalence and indifference. Second, we used effect sizes to evaluate incremental validity and adjusted for heteroskedasticity, providing confidence that our tests did not over- or underestimate the PN–SMD's utility or validity. Third, the CSI, the bipolar semantic differential, and the PANQIMS each provide a high bar for demonstrating incremental validity, thus making the substantial contribution of the PN–SMD over each of them rather impressive. This is especially so for the CSI, which represents the most informative items from well-validated and frequently used satisfaction measures.

Any interpretation of the present findings, however, should be tempered by the following study limitations. First, the extent to which the PN–SMD is incrementally valid across other criteria, or different operations of the same criteria, is uncertain. Moreover, we relied solely on self-report data as opposed to employing a multimethod approach to construct validation. Second, both samples predominantly comprised women. Although demographic characteristics were controlled for or separately examined, additional research is needed to ensure that the current findings replicate in other samples. Third, the present data comprised reports from only one dyad member; it is, therefore, unknown whether the current findings are robust when using couples as the unit of analysis. Fourth, the composition of our samples and the data they provided may somehow reflect the circumstances under which they were recruited (e.g., need for extra credit in psychology courses), the methods of recruitment (e.g., Internet advertisements), or some other aspect of the study or its design. Perhaps most notably, our participants were drawn from marital websites or college classrooms, which may have populated our samples with relatively satisfied individuals. However, this likely would make positive and negative attitudes more difficult to differentiate and, therefore, highlights the need for replication but does not undermine the thrust of our findings. Last, we assumed that the best semantic differential items to include on the PN–SMD were the ones that provided the most information in their one-dimensional format. It is plausible that using other, currently overlooked semantic anchors would provide for an incrementally more valid assessment of positive and negative attitudes toward the relationship. However, if we used alternative semantic anchors, it could be argued that any incremental validity over Funk and Rogge's (2007) bipolar item set was due to differences in item content, as opposed to the two-dimensional structure of our measure. Conversely, reformatting alternative items into a new bipolar scale—and then using this as our one-dimensional comparison index—would yield a different problem. Specifically, the most valid two-dimensional items may not serve as an adequate one-dimensional relationship satisfaction comparison measure in their bipolar format, so demonstrating incremental validity in that case would be less impressive. Taken together, our approach allowed for the best comparison of the one- and two-dimensional assessment models. However, as alternative positive and negative items may ultimately prove more useful, subsequent research can and should weigh their potential value against the currently validated PN–SMD.

Conclusion

The findings from the present studies support the two-dimensional conceptualization of relationship satisfaction. Furthermore, the

PN–SMD emerged as an internally consistent, two-factor measure that shows promise at providing an incrementally valid assessment of this construct. We hope that these data will stimulate research on its continued application and extension to more heterogeneous samples and alternate methods for assessing relationship functioning. Furthermore, the present findings are consistent with the larger corpus of attitude research demonstrating a two-factor model for evaluative judgments, which may allow relationship researchers to draw from this rich literature on the functional properties and behavioral correlates of positive and negative attitudes more generally.

References

Abakoumkin, G., Stroebe, W., & Stroebe, M. (2010). Does relationship quality moderate the impact of marital bereavement on depressive symptoms? *Journal of Social and Clinical Psychology*, *29*, 510–526.

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

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *A handbook of social psychology* (pp. 798–844). Worcester, MA: Clark University Press.

Bowman, M. L. (1990). Coping efforts and marital satisfaction: Measuring marital coping and its correlates. *Journal of Marriage and the Family*, *52*, 463–474.

Bradbury, T. N., Fincham, F. D., & Beach, S. R. H. (2000). Research on the nature and determinants of marital satisfaction: A decade in review. *Journal of Marriage and Family*, *62*, 964–980.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Taylor & Francis.

Cacioppo, J. T., Gardner, W. L., & Bernston, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychological Review*, *1*, 3–23.

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*, 189–225.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Conner, M., & Sparks, P. (2002). Ambivalence and attitudes. *European Review of Social Psychology*, *12*, 37–70.

Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics and Data Analysis*, *45*, 215–233.

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2006). Testing and probing interactions in hierarchical linear growth models. In C. S. Bergeman & S. M. Boker (Eds.), *The Notre Dame series on quantitative methodology: Vol. 1. Methodological issues in aging research* (pp. 99–129). Mahwah, NJ: Erlbaum.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Questionnaire.* London, UK: Hodder & Stoughton.

Fincham, F. D., Beach, S. R. H., & Kemp-Fincham, S. I. (1997). Marital quality: A new theoretical perspective. In R. J. Sternberg & M. Hojjat (Eds.), *Satisfaction in close relationships* (pp. 275–304). New York, NY: Guilford.

Fincham, F. D., & Bradbury, T. N. (1987). The assessment of marital quality: A reevaluation. *Journal of Marriage and the Family*, *49*, 797–809.

Fincham, F. D., & Bradbury, T. N. (1992). Assessing attributions in marriage: The relationship attribution measure. *Journal of Personality and Social Psychology*, *64*, 442–452.

Fincham, F. D., & Linfield, K. J. (1997). A new look at marital quality: Can spouses feel positive and negative about their marriage? *Journal of Family Psychology*, *11*, 489–502.

Fincham, F. D., & Rogge, R. D. (2010). Understanding relationship quality: Theoretical challenges and new tools for assessment. *Journal of Family Theory and Review*, *2*, 227–242.

Frame, L. E., Mattson, R. E., & Johnson, M. D. (2009). Predicting the success or failure of marital relationships. In H. T. Reis & S. K., Sprecher (Eds.), *Encyclopedia of human relationships.* (Vol. 3, pp. 1275–1279). Thousand Oaks, CA: Sage.

Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology*, *21*, 572–583.

Gable, S. L., Gonzaga, G. C., & Strachman, A. (2006). Will you be there for me when things go right? Supportive responses to positive event disclosures. *Journal of Personality and Social Psychology*, *91*, 904–917.

Gable, S. L., Reis, H. T., Impett, E. A., & Asher, E. R. (2004). What do you do when things go right? The intrapersonal and interpersonal benefits of sharing positive events. *Journal of Personality and Social Psychology*, *87*, 228–245.

Games, P. A., & Howell, J. F. (1976). Pairwise multiple comparison procedures with unequal n's and/or variances. *Journal of Educational Statistics*, *1*, 113–125.

Guilford, R., & Bengston, V. (1979). Measuring marital satisfaction in three generations: Positive and negative dimensions. *Journal of Marriage and the Family*, *41*, 387–398.

Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, *39*, 709–722.

Haynes, S. N., & Lench, H. C. (2003). Incremental validity of new clinical assessment measures. *Psychological Assessment*, *15*, 456–466.

Hudson, W., Harrison, D., & Crosscup, P. (1981). A short-form scale to measure sexual discord in dyadic relationships. *Journal of Sex Research*, *17*, 157–174.

Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment*, *15*, 1040–3590.

Huston, T. L., Caughlin, J. P., Houts, R. M., Smith, S. E., & George, L. J. (2001). The connubial crucible: Newlywed years as predictors of marital delight, distress, and divorce. *Journal of Personality and Social Psychology*, *80*, 237–252.

Jacobson, N. S. (1985). The role of observational measurement in behavior therapy outcome research. *Behavioral Assessment*, *7*, 297–308.

Johnson, D. R., White, L. K., Edwards, J. N., & Booth, A. (1986). Dimensions of marital quality: Toward methodological and conceptual refinement. *Journal of Family Issues*, *7*, 31–49.

Kaplan, K. J. (1972). On the ambivalence–indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique. *Psychological Bulletin*, *77*, 361–372.

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, methods, and research. *Psychological Bulletin*, *118*, 3–34.

Karney, B. R., & Bradbury, T. N. (1997). Neuroticism, marital interaction and the trajectory of marital satisfaction. *Journal of Personality and Social Psychology*, *78*, 295–309.

Kessler, R. C., Kendler, K. S., Heath, A., Neale, M. C., & Eaves, L. J. (1992). Social support, depressed mood, and adjustment to stress: A genetic epidemiologic investigation. *Journal of Personality and Social Psychology*, *62*, 257–272.

Klopfer, F. J., & Madden, T. M. (1980). The middlemost choice on attitude items: Ambivalence, neutrality, or uncertainty? *Personality and Social Psychology Bulletin*, *6*, 97–101.

Locke, H. J., & Wallace, K. M. (1959). Short marital-adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living*, *21*, 251–255.

Mattson, R. E., Paldino, D., & Johnson, M. D. (2007). The increased construct validity and utility of assessing relationship quality using separate positive and negative dimensions. *Psychological Assessment*, *19*, 146–151.

Menchaca, D., & Dehle, C. (2005). Marital quality and physiological arousal: How do I love thee? Let my heartbeat count the ways. *American Journal of Family Therapy*, *33*, 117–130.

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

Myers, J. L., & Well, A. D. (1995). *Research design and statistical analysis.* Mahwah, NJ: Erlbaum.

Norton, R. (1983). Measuring marital quality: A critical look at the dependent variable. *Journal of Marriage and the Family*, *45*, 141–151.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, *12*, 413–417.

Orden, S. R., & Bradburn, N. M. (1968). Dimensions of marriage happiness. *American Journal of Sociology*, *73*, 715–731.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana: University of Illinois Press.

Rodriguez, A., & Rogge, R. D. (2012). *When you war with your partner, everyone loses: The impact of negative relationship behavior on individual and dyadic well-being*. Manuscript in preparation.

Rusbult, C. E., Martz, J. M., & Agnew, C. R. (1998). The investment model scale: measuring commitment level, satisfaction level, quality of alternatives, and investment size. *Personal Relationships*, *5*, 357–391.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York, NY: Oxford University Press.

Snijders, T. A. B., & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods and Research*, *22*, 342–362.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*, *38*, 15–28.

Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistics* (5th ed.). Needham Heights, MA: Allyn & Bacon.

Watson, D., & Clark, L. A. (1991). *The mood and anxiety symptom questionnaire.* Unpublished manuscript, University of Iowa, Iowa City.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

Weiss, R. L. (2005). A critical view of marital satisfaction. In W. M. Pinsof & J. L. Lebow (Eds.), *Family psychology: The art of science.* New York, NY: Oxford University Press.

Appendix

Considering only the *positive qualities* of your relationship and *ignoring the negative ones*, evaluate your relationship on the following qualities:

My relationship is...

	Not at all 0	A tiny bit 1	A little 2	Somewhat 3	Mostly 4	Very 5	Extremely 6	Completely 7
Interesting	O	O	O	O	O	O	O	O
Full	O	O	O	O	O	O	O	O
Sturdy	O	O	O	O	O	O	O	O
Enjoyable	O	O	O	O	O	O	O	O
Good	O	O	O	O	O	O	O	O
Friendly	O	O	O	O	O	O	O	O
Hopeful	O	O	O	O	O	O	O	O

Considering only the *negative qualities* of your relationship and *ignoring the positive ones*, evaluate your relationship on the following qualities:

My relationship is...

	Not at all 0	A tiny bit 1	A little 2	Somewhat 3	Mostly 4	Very 5	Extremely 6	Completely 7
Bad	O	O	O	O	O	O	O	O
Lonely	O	O	O	O	O	O	O	O
Discouraging	O	O	O	O	O	O	O	O
Boring	O	O	O	O	O	O	O	O
Empty	O	O	O	O	O	O	O	O
Fragile	O	O	O	O	O	O	O	O
Miserable	O	O	O	O	O	O	O	O