Published in final edited form as:

Int J Cogn Ther. 2017 ; 10(1): 17–33. doi:10.1521/ijct_2016_09_18.

The Measurement of Cognitive Schemas: Validation of the Psychological Distance Scaling Task in a Community Mental Health Sample

Caroline Diehl¹, Seohyun Yin¹, Hannah Markell², Robert Gallop³, Mary Beth Connolly Gibbons¹, and Paul Crits-Christoph¹

¹Department of Psychiatry, University of Pennsylvania, Philadelphia, PA

²Department of Psychology, George Mason University, Fairfax, VA

³Department of Mathematics, West Chester University, West Chester, PA

Abstract

Objective—The purpose of this study was to examine the validity of the Psychological Distance Scaling Task (PDST), a measure of cognitive schema organization, in a community mental health setting. We also compared validity among African Americans and Caucasians.

Method—In order to accommodate participants with low education levels, 26 out of 80 PDST word stimuli were replaced with similar words at a lower reading level. A sample of 466 (42% African American; 50% Caucasian; 8% other) community patients with major depressive disorder completed the PDST and a variety of depressive symptom measures.

Results—The modified PDST demonstrated acceptable validity within all subscales. Validity coefficients resembled those reported in prior studies and were similar within minority and non-minority subsamples.

Conclusions—The modified PDST appears to be a valid measure of schema organization in a low-income, racially diverse population seeking treatment for depression at community clinics.

A major challenge in developing innovative diagnostic and treatment approaches for clinical depression is the identification of reliable vulnerability markers – that is, cognitive and psychosocial characteristics that predate or outlast depressive episodes and thus cannot be viewed as simply the byproducts of full-blown illness (Barnett & Gotlib, 1988). Initially, research suggested that cognitive variables such as negative biases, irrational beliefs, and skewed internal and external attributions are equally undetectable among healthy controls, formerly depressed individuals, and non-depressed individuals who later go on to develop depression, implying that depressive cognitions serve as transient episodic markers rather than premorbid risk factors (Dobson & Shaw, 1987; Dohr, Rush & Bernstein, 1989; Hammen, Miklowitz & Dyck, 1986; Lewinsohn et al., 1981). However, numerous priming studies have indicated that stress or negative mood induction can activate depressive cognitions among vulnerable individuals who display no current depressive symptoms

(Ingram & Siegle, 2009; Just, Abramson & Alloy, 2001; Scher, Ingram & Segal, 2005). It remains unclear whether those vulnerable to depression possess stable cognitive trait markers that can be measured in the absence of stress, affective priming, or active depression-related symptoms (Dozois & Dobson, 2001a).

The Psychological Distance Scaling Task (PDST), a computerized self-report measure, was developed by Dozois and Dobson (2001a, 2001b) to evaluate the internal structure of the self-schema - a set of self-referential cognitive heuristics that inform behavior rationalization and processing of self-relevant information (Beck, 1967; Dozois & Beck, 2008; Fiske & Linville, 1980; Markus, 1977). It had been theorized that schematic *structure* – defined as the organization and interconnectedness of positive and negative self-representations – could be assessed independently of schematic content and products (Ingram, Miranda & Segal, 1998). Dozois and Dobson (2001a) elaborated on this theory, proposing that while stress is a necessary catalyst for depressotypic schematic processing among vulnerable but asymptomatic individuals (Beck, 1967; Beck, Rush, Shaw, & Emery, 1979; Dozois & Beck, 2008), the internal organization of the depressive self-schema may be measurable regardless of affective state. In accordance with this hypothesis, depressotypic schematic structure – namely, highly interconnected negative content and highly diffuse positive content – appears to persist in fully remitted individuals who match healthy controls on other depressotypic cognitive indices (Dozois, 2007; Dozois & Dobson, 2001a). In addition, depressotypic schematic organization interacts with environmental stress to predict depression onset and symptomatic worsening (Hammen et al., 1985; Seeds & Dozois, 2010). Taken together, evidence from prospective and remission studies suggests that depressotypic schematic structure may serve as a stable, mood-independent trait marker for clinical depression (Dozois, 2007; Dozois & Dobson, 2001b; Dozois & Frewen, 2006; Lumley et al., 2012). It is important to note that depressotypic schematic structure is likely only one of multiple cognitive vulnerability markers for depression. However, given that schema organization may be measurable irrespective of mood state, it is of particular clinical interest.

Cognitive schema organization may also be a meaningful variable for understanding recovery over the course of cognitive therapy. A randomized trial comparing pharmacotherapy to joint treatment with cognitive therapy and pharmacotherapy found that both treatments successfully alleviated cognitive symptoms of depression, but only the joint treatment was associated with a measurable reversal of depressotypic schematic structure (Dozois et al., 2009). However, a more recent trial comparing cognitive behavioral therapy and pharmacotherapy found that both treatments elicited a significant reorganization of the depressive self-schema (Quilty et al., 2014). These contradictory results may be explained in part by differences in baseline illness severity; it is possible that pharmacotherapy is more capable of restructuring the depressive self-schema in moderate depression than in severe depression (Dozois & Quilty, 2013). However, further investigation is needed to clarify the nuances of this potential effect, as well as the broader relevance of cognitive structural change to long-term treatment efficacy (Garratt et al., 2007). Because the PDST is unique in its ability to evaluate enduring, mood-independent aspects of the depressive self-schema, it may serve as a vital tool for refining therapeutic strategies.

Before the PDST can be broadly applied, it must be validated in a variety of clinical contexts. The assessment has demonstrated high reliability thus far (Dozois, 2002; Dozois & Dobson, 2001b), but it was initially validated in an all-female undergraduate sample (Dozois, 2002) and has since been applied primarily in non-minority and/or highly educated populations. Specifically, the percentages of minority/African American participants in previous PDST studies were: Dozois (2007): 3%; Dozois et al. (2009): 2%; Quilty et al. (2014): 5.5% (African American). The validity of the PDST in African American/minority, low-income, and low-education samples has not been examined to date. Given the ongoing dissemination of cognitive therapy to community mental health settings that serve high proportions of low-income, low-education, and minority consumers (Creed et al., 2014), it is important to examine whether potential measures of constructs relevant to cognitive therapy perform with adequate psychometric properties in these settings and populations.

The purpose of this study was to assess the validity of the PDST in a low-income, racially diverse sample of male and female patients with depressive symptoms receiving treatment at a community mental health center. A second purpose was to compare the validity of the PDST among African Americans and Caucasians receiving treatment in this setting. Coefficients obtained in this sample were compared to those reported in prior studies in order to evaluate the validity of the PDST across a broad demographic range. To assess convergent validity, we evaluated the relation of the PDST to other measures of depressive cognitions, including measures of dysfunctional attitudes and compensatory skills. Although these cognitive measures assess somewhat different constructs than the PDST, it was hypothesized that patients whose PDST scores evidenced a depressogenic schematic structure would also demonstrate, to a moderate degree, depressotypic dysfunctional attitudes and negative compensatory skills. Given that PDST scores have been found to be associated with depressive symptom severity in previous studies (Dozois, 2002), we also hypothesized that PDST scores would be moderately correlated with depressive symptom measures, as well as a general measure of psychiatric symptoms. To examine the hypothesis that the PDST measures a construct relatively specific to the cognitive model of depression, we examined discriminant validity between the PDST and a measure designed from a psychodynamic perspective. This psychodynamic measure assessed self-understanding of interpersonal patterns, which is viewed as the central goal of certain psychodynamic approaches (e.g., Luborsky, 1984) but is distinct from the cognitive constructs (i.e., schematic organization) that the PDST purports to measure. Thus, we hypothesized that PDST scores would be unrelated to performance on a measure of interpersonal self-understanding. We also evaluated the extent to which PDST scores changed over the course of treatment for depression. Because the PDST is thought to measure a relatively stable vulnerability marker, we hypothesized that only modest changes would be evident over time.

Method

Participants

Participants were recruited at a community mental health center with a large outpatient clinic. All participants were screened for and/or participated in an ongoing large-scale effectiveness study being conducted at the center (Connolly Gibbons et al., 2014). Inclusion

criteria were as follows: 1) score of 11 or higher on the Quick Inventory of Depressive Symptoms (QIDS), indicating moderate to severe depressive symptoms; 2) 18–65 years of age. Participants were excluded if they reported any of the following: 1) acute medical problem requiring immediate inpatient treatment; 2) current substance abuse or dependence requiring immediate referral to a substance abuse program; 3) suicidal gesture within the past 3 months or significant suicidality requiring immediate medical referral. Informed consent was obtained from all individual participants included in the study.

Table 1 lists demographic and clinical characteristics of the patient sample ($N=466$). The sample was about 50% minority (primarily African American) and had substantially lower levels of education and employment than participants in previous studies involving the PDST. In addition, well over half (63.5%) of the current sample reported a household income of less than \$10,000 per year.

Measures

All measures were administered at baseline and, for patients enrolled in the effectiveness study, at months 1, 2, and 5 following baseline. Due to the restricted range of baseline depressive symptoms that resulted from recruiting patients based on depression severity, concurrent validity analyses focused on the last assessment obtained from each patient, regardless of when in the treatment process it occurred. Use of these endpoint assessments for all patients enrolled in the effectiveness study ($N=320$) allowed for adequate variability on depressive symptom measures that were used to examine the concurrent validity of the PDST. Baseline values were used for an additional 146 patients who did not receive a diagnosis of major depressive disorder (and therefore were excluded from the effectiveness study) but still had clinically meaningful levels of depressive symptoms. Sensitivity of the PDST to clinical change was assessed using all available evaluations (baseline and months 1, 2, and 5) for all patients with at least one post-baseline assessment.

The Hamilton Depression Inventory – 17 item (HAM-D; Hamilton, 1960)—The HAM-D is a widely used, clinician-administered inventory for evaluating the severity of depressive symptoms. In this study, the 17-item version of the HAM-D was administered by trained clinical evaluators using the Structured Interview Guide developed by Williams (1988), which has demonstrated good inter-judge reliability for a test-retest assessment of the composite HAM-D score (intraclass correlation coefficient = .81) (Williams, 1988). The clinical evaluators were nine advanced graduate students in clinical psychology who were blind to treatment condition and study hypotheses. They were trained and supervised by a senior clinician with expertise in delivering the HAM-D. The trainer provided written feedback to clinical evaluators based on a random review of 10% of audiotaped interviews and conducted a monthly group conference call with all clinical evaluators to maintain reliability. The internal consistency (Cronbach's alpha) of the HAM-D in the current sample was .76.

The Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996)—The BDI-II is a 21-item self-report measure that surveys depressive symptoms on a 4-point scale, with a focus on cognitions. It has demonstrated good internal consistency (Cronbach's $\alpha = .$

91; Beck, Steer, Ball, & Ranieri, 1996), as well as adequate construct validity with clinical ratings of depression ($r = .72$) (Beck et al., 1996). Reliability and validity of the BDI-II have been established in low-income and minority samples (Grothe et al., 2005; Joe et al., 2008). The BDI-II was found to have excellent internal consistency in the current study (Cronbach's $\alpha = .94$).

The Dysfunctional Attitudes Scale (DAS; Weissman & Beck, 1978)—The DAS (Form A) has been used extensively in studies of cognitive therapy outcomes (Dozois, Covin, & Brinker, 2003) and exhibits excellent psychometric characteristics (Nezu, Ronan, Meadows, & McClure, 2000). It consists of 40 attitudinal statements designed to represent depressotypic “underlying assumptions” (Beck et al., 1979), which consumers are asked to rate on a 7-point Likert scale according to their level of agreement. Several studies have associated change in DAS scores with alleviation of depressive symptoms among patients receiving CT for MDD (see review by Garratt et al., 2007). We found excellent internal consistency for this instrument in the present study (Cronbach's $\alpha = .94$).

The Self-Understanding of Interpersonal Patterns Scale-Revised (SUIP-R; Connolly et al., 1999)—The SUIP-R is a 28-item self-report inventory designed to measure the level at which an individual understands his or her own impairing relationship conflicts. This construct has been identified as a key mechanism of symptomatic change in some forms of psychodynamic psychotherapy (Gibbons et al., 2009). Consumers are first presented with a list of interpersonal patterns and asked to designate those that they currently experience in their own relationships. Consumers then rate each relevant pattern on a 6-point self-understanding scale, with the lowest rating indicating surface-level acknowledgment and the highest indicating a more complex understanding of personal contributions to the pattern and its development over time. A sample of 282 consumers demonstrated a high split-half internal consistency reliability on the SUIP-R (Cronbach's $\alpha = .92$), as well as strong discriminant validity, with self-understanding scores at baseline correlating less than $r = .10$ with measures of depressive symptoms, interpersonal distress, and quality of life (Gibbons et al., 2009). An increase in self-understanding as measured by the SUIP-R has been associated with a decrease in depressive symptoms and may be more likely to result from supportive-expressive psychodynamic therapy than from cognitive therapy (Gibbons et al., 2009). In the current sample, a split-half internal reliability coefficient of .92 was evident for the SUIP-R.

The Revised Behavior and Symptom Identification Scale-24 Item (BASIS-24, also known as BASIS-R; Eisen et al., 2004)—The BASIS-24 assesses an individual's impressions of his or her own behavioral health in six different categories – depression/functioning, interpersonal relationships, self-harm, emotional lability, psychosis, and substance abuse – using a 24-item self-report questionnaire. It has previously been validated in diverse, large-scale patient samples (Cameron et al., 2007), as well as specifically within Caucasian, African American, and Latino patient groups (Eisen et al., 2006), and has demonstrated adequate psychometric characteristics on all subscales. The depression/functioning subscale, which is most relevant to the present study, has exhibited sufficient internal consistency (Cronbach's α ranging from .87 to .90, Cameron et al., 2007;

Cronbach's α ranging from .85 to .91, Eisen et al., 2006), as well as high validity and sensitivity to symptomatic change (Eisen et al., 2006). In the current sample, the internal consistency of the BASIS-24 was good (Cronbach's $\alpha = .88$).

The Ways of Responding-Community Version (WOR-COMM; Scott et al., 2011)

—The WOR-COMM was adapted from the Ways of Responding (WOR) questionnaire originally developed by Barber & DeRubeis (1992), which assesses an individual's engagement of positive and negative compensatory skills. The WOR-COMM was designed to more accurately measure these skills among patients receiving behavioral health care in the community by presenting depressotypic scenarios more relevant to this population. The assessment involves reading a series of mood induction scenarios and reporting thoughts and hypothetical behaviors associated with each. The reported thoughts and behaviors are rated by a clinician based on the degree to which they represent positive compensatory strategies (scored on a positive subscale, WOR-COMM+) and negative/depressotypic responses (scored on a negative subscale, WOR-COMM-). The WOR-COMM has demonstrated acceptable internal consistency in both undergraduate samples (Cronbach's $\alpha = .80$, WOR-COMM+ and .82, WOR-COMM-) and clinical populations (Cronbach's $\alpha = .72$, WOR-COMM+ and .79, WOR-COMM-), as well as moderate convergent validity with the BDI-II and DAS in clinical populations (Connolly Gibbons et al., 2014). We found adequate internal consistency for both the positive (Cronbach's $\alpha = .83$) and negative (Cronbach's $\alpha = .81$) subscales in the present sample.

The Psychological Distance Scaling Task (PDST; Dozois & Dobson, 2001a)—In

this computerized task, participants are shown a series of 80 adjectives – 40 related to interpersonal experiences and 40 related to personal achievement. In each of these two categories, half of the adjectives are traits typically regarded as positive, and half are typically regarded as negative, resulting in four subscales: interpersonal-positive (IP+), interpersonal-negative (IP–), achievement-positive (A+), and achievement-negative (A–). Adjectives appear one at a time in the center of a square grid, which is quadrisected by two axes. The horizontal axis is labeled “Not at all like me” on the left of the grid and “Very much like me” on the right; the vertical axis is labeled “Very positive” at the top of the grid and “Very negative” at the bottom. As each adjective is presented, participants are instructed to indicate the point on the screen that best characterizes the word's self-relevance (on the x-axis) and valence (on the y-axis). The x- and y-coordinates of each adjective's placement are recorded. A distance score is then calculated for each adjective based on its average spatial distance from all other adjectives in its subscale that the participant judges as self-relevant – i.e., all adjectives in that subscale that have been assigned a positive x-coordinate (Dozois & Dobson, 2001b). Using the formula provided by Dozois and Dobson (2001b), these distance scores are squared and summed for each of the four subscales (IP+, IP–, A+, A–). The sum of squared distances is then divided by the total possible distance (i.e., non-self-relevant adjectives are excluded), and the square root of the resultant value is obtained. A smaller average interstimulus distance indicates a tightly interconnected self-schema, where self-referent beliefs of a certain valence are closely linked; a larger average interstimulus distance signifies a loosely connected self-schema, where similarly-valenced beliefs are less coherently organized (Dozois & Dobson, 2001b). If a participant categorizes fewer than two

adjectives as self-relevant within a given subscale, then an average distance score cannot be calculated for that subscale.

In this study, we modified some of the adjectives used in the PDST to increase the assessment's accessibility within a diverse population of community patients. Specifically, out of a total of 80 adjectives, we identified 26 that would be classified above an eighth grade reading level and replaced them with similar words that did not exceed this level, following word evaluation procedures used by Dozois (1999). Our research team generated a list of 31 potential alternative words, to which we added another 14 words that met our reading level criteria and had already been judged as adequate alternatives by Dozois (1999), but were not included in the standard PDST. A team of 13 independent judges then rated the 80 words included in the standard PDST, plus our additional 45 test words, on their emotional intensity, imagery (the extent to which each word evoked a sensory experience), and valence (positive vs. negative) using a 7-point scale (1 = extremely low/negative, 7 = extremely high/positive). Judges were college graduates or current undergraduates recruited from outside the research team and were blind to the purpose of the word rating task. No specific training in the task was provided. Inter-judge reliability of the mean ratings across judges, as assessed by an intraclass correlation coefficient, was .83 for emotional intensity, .77 for imagery, and .96 for valence.

Our final list of 26 replacement words, selected based on judges' ratings, included 8 specified by Dozois (1999), plus 18 of the 31 generated by our team. When choosing among words that received similar ratings from the judges, we ruled out potentially problematic words based on consensus from team members experienced in administering other measures in the community mental health setting. In addition, words in corresponding categories (e.g., IP+ and IP-) were matched on frequency of use in the English language. Word lists for the four subscales were statistically equivalent on word length, average frequency of use in the English language, emotional intensity, and strength of imagery (Dozois, 1999; 2007). Valence ratings were used to confirm the selection of the correct numbers of positive and negative words to replace deleted positive and negative words. Independent group *t*-tests comparing the 26 replacement words to the 54 original words that were retained showed that on average, there were no significant differences in emotional intensity ($M(SD) = 4.1(.9)$ for new; $4.3(.8)$ for old; $t(78) = 1.3$, $p = .21$), though the new words were rated somewhat lower on imagery ($M(SD) = 4.2(1.1)$ for new; $4.6(.8)$ for old; $t(78) = 2.0$, $p = .047$). Given that the overall level of imagery was acceptable for replacement words (4.2 on a 7-point scale), and there was a need to balance multiple variables across the four subscales as well as between replacements and originals, these 26 new words were deemed acceptable.

In order to ensure that these vocabulary changes did not affect the psychometric characteristics of the PDST, we piloted our modified version in a sample of 100 undergraduate students. To evaluate the adequacy of our 80 chosen adjectives in this student sample, we calculated item-total correlations for new and old adjectives on the self-relevance and valence dimensions. In the actual scoring of the PDST, only distances between pairs of self-relevant adjectives are used. However, the extent to which each adjective hangs together with other adjectives on a scale in regard to capturing the full range of self-relevance for each subject is important for evaluation of adjectives. Within all four subscales, average

item-total correlations were no lower for the 26 new adjectives than for the 54 old adjectives, with the one exception of the self-relevance dimension within the A+ subscale, for which new adjectives demonstrated a slightly lower average item-total correlation (average $r = .49$) than old adjectives (average $r = .56$).

The distributions of the final PDST scores were found to be non-normal. To account for this, log transformations were applied, as was done in previous studies of the PDST (e.g., Dozois, 2007).

Results

Adequacy of New Words in Clinical Sample

We re-examined whether the new adjectives that we substituted into the PDST performed as well as the old adjectives within this clinical sample. As with the student sample, the new adjectives demonstrated good item-total correlations (equal or higher on average than the old adjectives) within all four subscales on the self-relevance and valence dimensions, with the exception of a slightly lower average item-total correlation for the new adjectives (.53) compared to the old adjectives (.56) on the A- self-relevance dimension.

Validity

To assess convergent and discriminant validity, we correlated PDST subscale scores with other clinical measures. Detailed results are listed in Table 2, along with correlations among all of the validity measures. Given the number of correlations examined, we declared statistical significance at the .005 level, though the size of the correlations is of primary importance. There were small to moderate significant correlations between all four subscale scores (IP+, IP-, A+, and A-) and all clinical measures except the SUIP-R and, in the case of the IP+ subscale, the WOR-COMM+ and WOR-COMM-. Correlations of the PDST IP- and A- subscales with the HAM-D were of a similar magnitude as correlations of the DAS with the HAM-D. We also correlated PDST scores from all 466 participants with the following seven demographic variables: age, minority status, gender, employment status, education level, relationship status, and income. Only one correlation reached the .005 significance level: decreased distance within the A+ subscale of the PDST was weakly but significantly associated with minority status ($r = -.17$, $p < .001$, $N = 435$) (members of minority groups had more tightly connected positive achievement responses).

Validity in African Americans vs. Caucasians

Finally, we compared the validity of the PDST within our African American subsample (N ranging from 105 to 192) – the largest minority subsample in our study population – to validity within the Caucasian majority subsample (N ranging from 145 to 232) (Table 3). To explore differential validity, we conducted regression analyses predicting PDST subscale scores from other measures and included cross-product terms (African American status by validity variable) after main effects to test for interactions. To enable the interpretation of any such interactions, regression coefficients within the African American and Caucasian subsamples are presented in Table 3.

There was little evidence of differential validity coefficients within these two racial groups. None of the cross-product terms involving African American status reached significance at the $p < .005$ level, suggesting a lack of significant difference in validity of the PDST for African Americans vs. Caucasians. However, the interaction term of African American status by WOR-COMM– for the IP– subscale was close to reaching significance ($t(272) = -2.50, p = .013$). Among African Americans, the presence of negative compensatory skills was more strongly associated with negative interpersonal self-schema than it was for Caucasians.

Change over Time

To examine the sensitivity of the PDST to clinical change in our community sample, we conducted hierarchical linear modeling analyses of the PDST scores available at baseline and months 1, 2, and 5. Because follow-up assessments were not always conducted precisely at the 1-, 2-, or 5-month mark, sensitivity analyses used days from baseline as the measure of time. Table 4 provides mean (SD) PDST scores at each assessment. For the two positive PDST subscales, there was a significant decrease in spatial distance over time (baseline to month 5), and for the two negative subscales, there was a significant increase in distance. However, no significant change occurred between baseline and month 1 (paired t-tests: IP+: $t(150) = 1.69, p = .09$; IP–: $t(149) = .83, p = .41$; A+: $t(132) = .87, p = .39$; A–: $t(139) = .20, p = .85$). Cohen's d effect sizes for the change from baseline to month 5 were modest, ranging from .18 to .47 (Table 4). In contrast, the baseline to month 5 effect size for change in the BDI-II was .70.

Discussion

We found that a version of the PDST calibrated to an eighth grade reading level demonstrated sufficient internal reliability in a large-scale, racially diverse, low-income sample of community patients with depressive symptoms. Furthermore, the modified PDST demonstrated convergent validity with several symptom and dysfunctional cognition-related measures – specifically, the HAM-D, BDI-II, DAS, BASIS, and WOR-SR, as well as the negative subscale of the WOR-COMM.

The sample recruited for this study demonstrated substantially greater racial, ethnic, and socioeconomic diversity than samples recruited for previous studies involving the PDST. Whereas previous samples consisted of currently enrolled undergraduates (Dozois, 2002; Seeds & Dozois, 2010), adult participants with average education levels ranging from 13.8 years to 15.6 years (Dozois, 2007; Dozois et al., 2009; Dozois & Dobson, 2001a, 2001b), or samples whose educational attainment was not reported in detail (Dozois & Frewen, 2006), the majority of the current sample had received only a high school education (12 years) or less. Similarly, employment rate was substantially lower in the current sample than in previous samples. The proportion of participants who were employed at the time of assessment was reported in three previous non-undergraduate samples and ranged from 46% (Dozois, 2007) to 64% (Dozois et al., 2009); conversely, only 15.8% of the current sample reported part- or full-time employment.

Despite these stark demographic differences, convergent validities of the PDST with the BDI-II and DAS resembled those reported by Dozois (2002). Correlation strength was slightly weaker in the current sample than in the Dozois (2002) sample, though these differences were not tested for statistical significance. For the negative summary scale, we report that $r = -.28$ with the DAS, and in the Dozois (2002) sample, $r = -.41$ with the DAS. In the present study, the positive summary scale correlated with the DAS at $r = .21$, and in the Dozois sample, $r = .35$. Similarly, for BDI-II scores, the positive summary scale showed a weaker ($r = .21$) correlation in the current sample compared to the Dozois sample ($r = .58$). It is possible that this discrepancy can be attributed to clinical differences between the two samples. Alternatively, changes made to the word content of the PDST in the present study may explain these differences.

The modified PDST also demonstrated satisfactory discriminant validity in our community sample. No PDST subscales correlated significantly with the SUIP-R. The SUIP-R has been validated as a measure of self-understanding of maladaptive relationship patterns, a construct that is known to mediate change in dynamic therapies (Gibbons et al., 2009). In contrast, the PDST was designed to measure cognitive schema organization, a central concept in cognitive therapy. The absence of any significant relationship between PDST and SUIP-R scores supports the inference that the PDST assesses constructs that are primarily relevant to a cognitive therapeutic model and less relevant to a psychodynamic model.

There were no definitive differences in the validity of the PDST for African American participants compared to Caucasians. For the negative interpersonal subscale, the interaction of race with WOR-COMM- score did not reach our .005 level of significance but had a p value of .013, suggesting the possibility that among African Americans, negative coping behaviors may have a greater negative bearing on the interpersonal self-concept than they do among whites. It is also possible that the sensitivity of the WOR-COMM negative subscale and/or the PDST interpersonal negative subscale is different for African Americans and whites. These possibilities should be explored in future research. However, the overall lack of significant interactions between race and validity measures suggests that on the whole, the validity of the PDST is likely similar for African Americans and Caucasians.

Both the interpersonal and achievement subscales of the PDST showed small to moderate correlations with depression-related measures. Previous studies have indicated that the structure of interpersonal schematic content – particularly negative interpersonal content – is more stable than the structure of achievement-related content and more likely to persist after depression symptoms have remitted (Dozois, 2007). This suggests that the interpersonal and achievement dimensions may play different roles in the initiation and maintenance of depression, despite having comparable relations with the severity of active depressive symptoms and dysfunctional cognitions. Further evidence is needed to clarify whether cognitive schematic structures that are more temporally stable also have superior predictive value with regard to depression-related cognitions and symptoms.

Analyses of change over time revealed that PDST scores showed no significant change from baseline to month 1, but that each subscale exhibited modest changes from baseline to month 5. These data suggest that the PDST was sensitive to clinical change in this

community sample of depressed patients. However, PDST scores demonstrated greater stability than did other symptom measures, which is consistent with the hypothesis that the PDST measures a relatively stable underlying construct that is most likely to change over an extended period of time.

The present study was limited by only including patients who had depressive symptoms. The Dozois (2002) study included three distinct clinical groups, labeled as nondysphoric, mildly dysphoric, and moderately to severely dysphoric, and only N=19 out of 78 (24.4%) of participants fell into the latter group, which had a mean BDI-II score of 27.6 ($SD = 7.0$) (Dozois, 2002). In contrast, all participants in the current study had some degree of depressive symptoms, leading to greater baseline uniformity and higher average symptom severity, as measured by the BDI-II ($M = 34.4$; $SD = 11.1$). In addition, PDST scoring conventions may have led to the exclusion of participants with the most and least extreme cognitive symptoms. Because PDST scores represent interstimulus distance, a composite score for a given PDST subscale can only be calculated if a participant rates at least two adjectives in that subscale as self-relevant (i.e., greater than 0 on the x-axis). If a participant categorizes fewer than two subscale adjectives as self-relevant, then the subscale score is recorded as missing, and the participant's ratings of adjectives within that subscale have no bearing on validity analyses. This scoring method can lead to the omission of participants with particularly mild symptoms, who may classify fewer than two negative adjectives as self-relevant, as well as participants with severe symptoms, who may not view any positive adjectives as self-relevant. In a severely depressed clinical sample such as the one recruited for the present study, the latter situation is particularly likely. This limitation may account for the lack of significant correlations between the A+ subscale and the HAM-D and BDI-II, as well as the weakness of the correlation between the positive summary score and the BDI-II compared with the correlation reported by Dozois (2002). However, establishing the reliability and validity of the PDST is an appropriate precursor for using this assessment to evaluate the role of schematic structural change in the process of cognitive therapy for MDD. In future reports, we plan to examine that role.

Conclusions

The PDST was found to exhibit acceptable internal reliability, as well as convergent and discriminant validity, in a diverse sample of low-income community mental health patients seeking treatment for depression. Validity coefficients were, for the most part, similar to those obtained previously in a primarily non-minority undergraduate sample (Dozois, 2002). The PDST may serve as a useful tool for measuring cognitive schema organization in community mental health centers and other settings with low-income, diverse populations.

Acknowledgments

This research was supported by National Institute of Mental Health Grant R01-MH092363. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Barber JP, DeRubeis RJ. The Ways of Responding: A Scale to Assess Compensatory Skills Taught in Cognitive Therapy. *Behavioral Assessment*. 1992; 14:93–115.
- Barnett PA, Gotlib IH. Psychosocial Functioning and Depression: Distinguishing Among Antecedents, Concomitants, and Consequences. *Psychological Bulletin*. 1988; 104(1):97–126. DOI: 10.1037/0033-2909.104.1.97 [PubMed: 3043529]
- Beck, AT. *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper & Row; 1967.
- Beck, AT., Rush, AJ., Shaw, BF., Emery, G. *Cognitive therapy of depression*. New York: Guilford Press; 1979.
- Beck AT, Steer RA, Ball R, Ranieri W. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment*. 1996; 67(3):588–597. DOI: 10.1207/s15327752jpa6703_13 [PubMed: 8991972]
- Beck, AT., Steer, RA., Brown, GK. *Manual for the Beck Depression Inventory–II*. San Antonio, TX: Psychological Corporation; 1996.
- Beck AT, Steer RA, Garbin MM. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*. 1988; 8(1):77–100. DOI: 10.1016/0272-7358(88)90050-5
- Blatt SJ. Levels of object representation in anaclitic and introjective depression. *The Psychoanalytic Study of the Child*. 1974; 29(10):7–157.
- Cameron IM, Cunningham L, Crawford JR, Eagles JM, Eisen SV, Lawton K, ... Hamilton RJ. Psychometric properties of the BASIS-24© (Behaviour and Symptom Identification Scale-Revised) Mental Health Outcome Measure. *International Journal of Psychiatry in Clinical Practice*. 2007; 11(1):36–43. DOI: 10.1080/13651500600885531 [PubMed: 24941274]
- Connolly MB, Crits-Christoph P, Shelton RC, Hollon S, Kurtz J, Barber JP, ... Thase ME. The reliability and validity of a measure of self-understanding of interpersonal patterns. *Journal of Counseling Psychology*. 1999; 46(4):472–482. DOI: 10.1037/0022-0167.46.4.472
- Connolly Gibbons MB, Gallop R, Barber JP, Temes C, Goldstein L, Stirman S, Crits-Christoph P. Unique and common mechanisms of change across cognitive and dynamic psychotherapies. *Journal of Consulting and Clinical Psychology*. 2009; 77(5):801–813. DOI: 10.1037/a0016596 [PubMed: 19803561]
- Connolly Gibbons, MB., Lee, JK., Mack, RA., Markell, HM., Crits-Christoph, P. The reliability and validity of community and self-report versions of the Ways of Responding Questionnaire. Poster presented at the 45th annual meeting of the Society for Psychotherapy Research; Copenhagen, Denmark. 2014; 2014 Jun.
- Connolly Gibbons MB, Mack R, Lee J, Gallop R, Thompson D, Burock D, Crits-Christoph P. Comparative effectiveness of cognitive and dynamic therapies for major depressive disorder in a community mental health setting: Study protocol for a randomized non-inferiority trial. *BMC Psychology*. 2014; 2(1):47.doi: 10.1186/s40359-014-0047-y [PubMed: 25566391]
- Creed TA, Stirman SW, Evans AC, Beck AT. A model for implementation of cognitive therapy in community mental health: The Beck Initiative. *The Behavior Therapist*. 2014; 37(3)
- Dobson KS, Shaw BF. Specificity and stability of self-referent encoding in clinical depression. *Journal of Abnormal Psychology*. 1987; 96(1):34–40. DOI: 10.1037/0021-843X.96.1.34 [PubMed: 3558947]
- Dohr KB, Rush AJ, Bernstein IH. Cognitive biases and depression. *Journal of Abnormal Psychology*. 1989; 98(3):263–267. DOI: 10.1037/0021-843X.98.3.263 [PubMed: 2788667]
- Dozois, DJ. Unpublished doctoral dissertation. The University of Calgary; Calgary, Alberta: 1999. Cognitive Organization and Information Processing in Clinical Depression: The Structure and Function of Sociotropic Schemata.
- Dozois DJA. Stability of negative self-structures: a longitudinal comparison of depressed, remitted, and nonpsychiatric controls. *Journal of Clinical Psychology*. 2007; 63(4):319–338. DOI: 10.1002/jclp.20349 [PubMed: 17279521]

- Dozois DJ. Cognitive Organization of Self-Schematic Content in Nondysphoric, Mildly Dysphoric, and Moderately–Severely Dysphoric Individuals. *Cognitive Therapy and Research*. 2002; 26:417–29.
- Dozois, DJA., Beck, AT. Cognitive schemas, beliefs and assumptions. In: Dobson, KS., Dozois, DJA., editors. *Risk factors in depression*. Oxford, England: Elsevier/Academic Press; 2008. p. 121–143.
- Dozois DJA, Bieling PJ, Patelis-Siotis I, Hoar L, Chudzik S, McCabe K, Westra HA. Changes in self-schema structure in cognitive therapy for major depressive disorder: a randomized clinical trial. *Journal of Consulting and Clinical Psychology*. 2009; 77(6):1078–1088. DOI: 10.1037/a0016886 [PubMed: 19968384]
- Dozois DJA, Covin R, Brinker JK. Normative data on cognitive measures of depression. *Journal of Consulting and Clinical Psychology*. 2003; 71(1):71–80. DOI: 10.1037/0022-006X.71.1.71 [PubMed: 12602427]
- Dozois DJA, Dobson KS. A Longitudinal Investigation of Information Processing and Cognitive Organization in Clinical Depression: Stability of Schematic Interconnectedness. *Journal of Consulting and Clinical Psychology*. 2001a; 69(6):914–925. DOI: 10.1037//0022-006X.69.6.914 [PubMed: 11777119]
- Dozois DJA, Dobson KS. Information processing and cognitive organization in unipolar depression: specificity and comorbidity issues. *Journal of Abnormal Psychology*. 2001b; 110(2):236–246. DOI: 10.1037/0021-843X.110.2.236 [PubMed: 11358018]
- Dozois DJA, Frewen PA. Specificity of cognitive structure in depression and social phobia: A comparison of interpersonal and achievement content. *Journal of Affective Disorders*. 2006; 90(2–3):101–109. DOI: 10.1016/j.jad.2005.09.008 [PubMed: 16343641]
- Dozois, DJA., Quilty, LC. Treatment changes in the depressive self-schema. *Psychological Science Agenda*. Jun. 2013 Retrieved from <http://www.apa.org/science/about/psa/2013/06/depressive-self-schema.aspx>
- Eisen SV, Gerena M, Ranganathan G, Esch D, Idieulla T. Reliability and validity of the BASIS-24 Mental Health Survey for Whites, African Americans, and Latinos. *Journal of Behavioral Health Services & Research*. 2006; 33(3):304–323. DOI: 10.1007/s11414-006-9025-3 [PubMed: 16752108]
- Eisen SV, Normand SL, Belanger AJ, Spiro A III, Esch D. The Revised Behavior and Symptom Identification Scale (BASIS-R): reliability and validity. *Medical Care*. 2004; 42(12):1230–1241. DOI: 10.1097/00005650-200412000-00010 [PubMed: 15550803]
- First, MB., Spitzer, RL., Gibbon, M., Williams, JBW. *Structured clinical interview for Axis I DSM-IV disorders (Version 2.0)*. Biometrics Research Department, New York State Psychiatric Institute; New York, NY: 1996.
- Fiske ST, Linville PW. What Does the Schema Concept Buy Us? *Personality and Social Psychology Bulletin*. 1980; 6(4):543–557. DOI: 10.1177/014616728064006
- Garratt G, Ingram RE, Rand KL, Sawalani G. Cognitive Processes in Cognitive Therapy: Evaluation of the Mechanisms of Change in the Treatment of Depression. *Clinical Psychology: Science and Practice*. 2007; 14(3):224–239. DOI: 10.1111/j.1468-2850.2007.00081.x
- Gloaguen V, Cottraux J, Cucherat M, Blackburn IM. A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders*. 1998; 49(1):59–72. DOI: 10.1016/S0165-0327(97)00199-7 [PubMed: 9574861]
- Grothe KB, Dutton GR, Jones GN, Bodenlos J, Ancona M, Brantley PJ. Validation of the Beck Depression Inventory-II in a low-income African American sample of medical outpatients. *Psychological Assessment*. 2005; 17(1):110–114. DOI: 10.1037/1040-3590.17.1.110 [PubMed: 15769232]
- Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry*. 1960; 23(1):56–62. DOI: 10.1136/jnnp.23.1.56
- Hammen C, Marks T, Mayol A, deMayo R. Depressive Self-Schemas, Life Stress, and Vulnerability to Depression. *Journal of Abnormal Psychology*. 1985; 94(3):308–319. DOI: 10.1037/0021-843X.94.3.308 [PubMed: 4031228]

- Hammen C, Miklowitz DJ, Dyck DG. Stability and severity parameters of depressive self-schema responding. *Journal of Social and Clinical Psychology*. 1986; 4(1):23–45. DOI: 10.1521/jscp.1986.4.1.23
- Ingram, RE., Miranda, J., Segal, ZV. *Cognitive vulnerability to depression*. New York: Guilford Press; 1998.
- Ingram, RE., Siegle, GJ. Methodological issues in the study of depression. In: Gotlib, IH., Hammen, CL., editors. *Handbook of depression*. 2. New York, NY: The Guilford Press; 2009. p. 69-92.
- Joe S, Woolley ME, Brown GK, Ghahramanlou-Holloway M, Beck AT. Psychometric properties of the Beck Depression Inventory–II in low-income, African American suicide attempters. *Journal of Personality Assessment*. 2008; 90(5):521–523. DOI: 10.1080/00223890802248919 [PubMed: 18704812]
- Just N, Abramson LY, Alloy LB. Remitted depression studies as tests of the cognitive vulnerability hypotheses of depression onset: A critique and conceptual analysis. *Clinical Psychology Review*. 2001; 21(1):63–83. DOI: 10.1016/S0272-7358(99)00035-5 [PubMed: 11148896]
- Lewinsohn PM, Steinmetz JL, Larson DW, Franklin J. Depression-Related Cognitions: Antecedent or Consequence? *Journal of Abnormal Psychology*. 1981; 90(3):213–219. DOI: 10.1037/0021-843X.90.3.213 [PubMed: 7288016]
- Luborsky, L. *Principles of psychoanalytic psychotherapy: A manual for supportive-expressive (SE) treatment*. New York, NY: Basic Books; 1984.
- Lumley MN, Dozois DJA, Hennig KH, Marsh A. Cognitive Organization, Perceptions of Parenting and Depression Symptoms in Early Adolescence. *Cognitive Therapy and Research*. 2012; 36(4): 300–310. DOI: 10.1007/s10608-011-9365-z
- Markus H. Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*. 1977; 35(2):63–78. DOI: 10.1037/0022-3514.35.2.63
- Nezu, AM., Ronan, GF., Meadows, EA., McClure, KS. *Practitioner's Guide to Empirically Based Measures of Depression*. New York, NY: Kluwer Academic/Plenum Publishers; 2000.
- Quilty LC, Dozois DJA, Lobo DSS, Ravindran LN, Bagby RM. Cognitive Structure and Processing During Cognitive Behavioral Therapy vs. Pharmacotherapy for Depression. *International Journal of Cognitive Therapy*. 2014; 7(3):235–250. DOI: 10.1521/ijct.2014.7.3.235
- Scher CD, Ingram RE, Segal ZV. Cognitive reactivity and vulnerability: Empirical evaluation of construct activation and cognitive diatheses in unipolar depression. *Clinical Psychology Review*. 2005; 25(4):487–510. DOI: 10.1016/j.cpr.2005.01.005 [PubMed: 15914266]
- Scott, K., Connolly Gibbons, MB., Schauble, LA., Thompson, SM., Heintz, LE., Hamilton, JL., Crits-Christoph, P. The development of community-friendly and self-report versions of the ways of responding questionnaire. Poster presented at the meeting of the Society for Psychotherapy Research, University of Bern; Switzerland. 2011 Jun.
- Seeds PM, Dozois DJA. Prospective Evaluation of a Cognitive Vulnerability-Stress Model for Depression: The Interaction of Schema Self-Structures and Negative Life Events. *Journal of Clinical Psychology*. 2010; 66(12):1307–1323. DOI: 10.1002/jclp.20723 [PubMed: 20715020]
- Segal ZV, Gemar M, Williams S. Differential cognitive response to a mood challenge following successful cognitive therapy or pharmacotherapy for unipolar depression. *Journal of Abnormal Psychology*. 1999; 108(1):3–10. DOI: 10.1037/0021-843X.108.1.3 [PubMed: 10066988]
- Segal ZV, Kennedy S, Gemar M, Hood K, Pedersen R, Buis T. Cognitive reactivity to sad mood provocation and the prediction of depressive relapse. *Archives of General Psychiatry*. 2006; 63(7): 749–755. DOI: 10.1001/archpsyc.63.7.749 [PubMed: 16818864]
- Weissman, AN., Beck, AT. Development and validation of the Dysfunctional Attitudes Scale: a preliminary investigation. *Proceedings of the Meeting of the American Educational Research Association*; Toronto, Ontario. 1978.
- Williams JBW. A Structured Interview Guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry*. 1988; 45(8):742–747. DOI: 10.1001/archpsyc.1988.01800320058007 [PubMed: 3395203]
- Zuroff DC, Blatt SJ, Sanislow CA III, Bondi CM, Pilkonis PA. Vulnerability to Depression: Reexamining State Dependence and Relative Stability. *Journal of Abnormal Psychology*. 1999; 108(1):76–89. DOI: 10.1037/0021-843X.108.1.76 [PubMed: 10066995]

Table 1

Patient Demographics and Clinical Characteristics

Demographic Variables (<i>N</i> =466)	<i>N</i> (%)
Gender (male)	122 (26.2%)
Race	
White	234 (50.2%)
Black/African American	194 (41.6%)
Other/Unknown	38 (8.2%)
Ethnicity (Hispanic)	19 (4.1%)
Marital Status (currently married)	46 (9.9%)
Education	
Less than 12 Years	95 (20.4%)
High School Diploma/GED	191 (41.0%)
Some College	116 (24.9%)
Associate's Degree	35 (7.5%)
Bachelor's Degree or beyond	29 (6.3%)
Employment (employed full or part time)	74 (15.8%)
Age, years <i>M</i> (<i>SD</i>)	36.7 (12.1)
Household Income	
0–\$10,000	296 (63.5%)
>\$10,000	122 (26.2%)
Clinical Severity Measures (417 <i>N</i> 463)	<i>M</i> (<i>SD</i>)
PDST Mean Scores at Baseline	
IP+	1.18 (.28)
IP–	1.56 (.36)
A+	1.29 (.35)
A–	1.48 (.39)
Depressive Symptoms Mean Scores	
HAM-D	17.6 (7.2)
BDI-II	29.7 (13.4)

Validity Coefficients (Pearson Correlations) for the Psychological Distance Scaling Task among Community Mental Health Center Patients with Major Depressive Disorder

Table 2

Validity Measures	PDST					Validity Measures				
	IP+	IP-	A+	A-	HAMD	BDI-II	DAS	SUIP-R	BASIS-24	WOR+
HAMD-D	.19***	-.23***	.17***	-.27***						
BDI-II	.26***	-.28*	.25***	-.39***	.63***					
DAS	.29***	-.32*	.23***	-.35***	.30***	.49***				
SUIP-R	.08	-.02	.02	-.04	.07	.14**	.04			
BASIS-24	.28***	-.23*	.27***	-.35***	.59***	.82***	.41***	.14*		
WOR+	-.06	.19*	-.17***	.18***	-.21***	-.25***	-.38***	.14*	-.29***	
WOR-	.15	-.22*	.23***	-.25***	.22***	.35***	.47***	.12*	.30***	-.34***

Note. PDST, Psychological Distance Scaling Task; IP+, interpersonal positive subscale; IP-, interpersonal negative subscale; A+, achievement positive subscale; A-, achievement negative subscale; HAMD, Hamilton Depression Inventory-17 Item; BDI-II, Beck Depression Inventory-II; DAS, Dysfunctional Attitudes Scale; SUIP-R, Self-Understanding of Interpersonal Patterns Scale-Revised (self-understanding scale); BASIS-24, depression subscale of the Behavior and Symptom Identification Scale-24 Item; WOR+ and WOR-, positive and negative subscales of the Ways of Responding-Community Version. Sample size for each correlation ranges from N=277 to N=460 due to missing data on some measures and the fact that the WOR was not scored for all patients.

* p < .05;

** p < .01;

*** p < .005.

Table 3

Validity of the PDST in African American vs. White Sub-samples

Validity Measures	PDST Subscales							
	IP+		IP-		A+		A-	
	AA	W	AA	W	AA	W	AA	W
HAM-D	2.51 (1.70)	7.39*** (1.74)	-4.87*** (1.34)	-4.19*** (1.33)	2.28 (1.74)	3.44* (1.34)	-3.41* (1.34)	-5.82*** (1.26)
BDI-II	8.11* (3.1)	16.94*** (3.26)	-8.74*** (2.46)	-10.92*** (2.41)	8.26** (3.11)	9.16*** (2.49)	-10.14*** (2.36)	-15.01*** (2.20)
DAS	20.92* (8.43)	56.89*** (9.06)	-25.40*** (6.87)	-40.86*** (6.80)	16.27 (8.26)	25.12*** (6.96)	-26.05*** (6.46)	-40.02*** (6.60)
SUIP-R	-.02 (.30)	.71* (.35)	-.29 (.24)	.16 (.27)	-.14 (.31)	.12 (.27)	.03 (.24)	-.26 (.26)
BASIS-24	.64*** (.20)	1.20*** (.22)	-.50*** (.16)	-.61*** (.17)	.64*** (.20)	.75*** (.17)	-.65*** (.16)	-.92*** (.15)
WOR+	.08 (.27)	-.54 (.31)	.29 (.19)	.72*** (.23)	-.29 (.24)	-.62** (.22)	.18 (.20)	.62*** (.20)
WOR-	.13 (.26)	.90** (.32)	-.68*** (.18)	-.32 (.25)	.30 (.24)	.68*** (.21)	-.54* (.21)	-.77*** (.22)

Note. Unstandardized regression coefficients are given, with standard errors in parentheses. PDST, Psychological Distance Scaling Task; IP+, interpersonal positive subscale; IP-, interpersonal negative subscale; A+, achievement positive subscale; A-, achievement negative subscale; HAM-D, Hamilton Depression Inventory-17 Item; BDI-II, Beck Depression Inventory-II; DAS, Dysfunctional Attitudes Scale; SUIP-R, Self-Understanding of Interpersonal Patterns Scale-Revised (self-understanding scale); BASIS-24, depression subscale of the Behavior and Symptom Identification Scale-24 Item; WOR+ and WOR-, positive and negative subscales of the Ways of Responding-Community Version. AA, African American; W, white. Sample size for the African American subsample ranges from N=105 to N=192; sample size for the white subsample ranges from N=145 to N=232. Regression analyses predicting log-transformed PDST subscale scores showed no significant interactions between race (AA vs. W) and score on validity measures at the $p < .005$ level. For regression beta coefficients:

* $p < .05$;

** $p < .01$;

*** $p < .005$.

Table 4

Change in PDST Subscale Scores over Time

PDST subscale	Mean (SD) at Each Evaluation				Time Effect		
	Baseline	Month 1	Month 2	Month 5	F (DF)	p	Effect Size (d): baseline to month 5
IP+	1.25 (.28)	1.23 (.30)	1.20 (.29)	1.17 (.27)	56.9 (1, 686)	< .001	.29
IP−	1.52 (.30)	1.49 (.34)	1.58 (.34)	1.66 (.36)	135.8 (1, 686)	< .001	.47
A+	1.39 (.36)	1.41 (.43)	1.33 (.35)	1.26 (.38)	7.97 (1, 641)	.005	.36
A−	1.45 (.34)	1.41 (.35)	1.47 (.38)	1.51 (.41)	75.3 (1, 628)	< .001	.18

Note. N ranges from 115 to 230 depending on measure and assessment point.