Title
Briefly Assessing Repetitive Thought Dimensions

Journal
Assessment, 23(5)

ISSN
1073-1911

Authors
Segerstrom, Suzanne C
Hardy, Jaime K
Evans, Daniel R
et al.

Publication Date
2016-10-01

DOI
10.1177/1073191115586458

Peer reviewed

Published in final edited form as:

Assessment. 2016 October ; 23(5): 614–623. doi:10.1177/1073191115586458.

Briefly Assessing Repetitive Thought Dimensions: Valence, Purpose, and Total

Suzanne C. Segerstrom[*],
Department of Psychology, University of Kentucky, Lexington, KY

Jaime K. Hardy,
Department of Psychology, University of Kentucky, Lexington, KY

Daniel R. Evans,
Department of Psychology, University of Kentucky, Lexington, KY

Ian A. Boggero,
Department of Psychology, University of Kentucky, Lexington, KY

Lynn E. Alden, and
Department of Psychology, University of British Columbia, Vancouver, BC

Annette L. Stanton
Departments of Psychology and Psychiatry/Biobehavioral Sciences, University of California, Los Angeles, CA

Abstract

Discrete forms of repetitive thought (RT) such as worry and reflection can be characterized along basic dimensions of valence (positive vs. negative) and purpose (searching vs. solving). In addition, people can be characterized as high or low in their tendency to engage in RT. This dimensional model has been demanding to assess, and a smaller number of items that could stand in for a large battery would make measurement more accessible. Using 4 samples (N = 1588), 8 items were identified that assess RT valence, purpose, and total in a circumplex model. Across these and other samples, the dimensions were adequately reliable and valid with regard to assessment via large RT battery, other measures of RT, and depressive symptoms. The accessibility of dimensional assessment of RT using this smaller number of items should facilitate work on questions about the qualities of RT that predict mental and physical health.

Keywords

repetitive thought; multidimensional; assessment; short form; diary

Repetitive thought (RT) is defined as thinking attentively, repetitively, and/or frequently about oneself and one's world (Segerstrom, Stanton, Alden, & Shortridge, 2003). It is an umbrella construct that has transdiagnostic significance for emotional and physical well-

[*]Corresponding author: Department of Psychology, University of Kentucky, 125 Kastle Hall, Lexington, KY 40506-0044, Phone 859-257-4549, FAX 859-323-1979, segerstrom@uky.edu.

being (Segerstrom et al., 2003; Segerstrom, Roach, Evans, Schipper, & Darville, 2010; Watkins, 2008) and encompasses a number of discrete constructs, including depressive rumination, reflection, planning, emotional processing, reminiscing, and anticipating. All of these constructs suitably fall under the umbrella of RT because they positively correlate with each other – even constructs such as reflection and worry that appear different from each other. Mean correlations among diverse RT questionnaires typically range from .20 – .40 (Evans & Segerstrom, 2011; Segerstrom et al., 2003; Segerstrom et al., 2010). Therefore, there are individual differences in how much people engage in repetitive thought of all kinds, and a sum of scores on RT questionnaires (i.e., total RT) can represent the extent of this tendency to engage in RT.

This quantitative dimension of RT, however, fails to capture qualitative differences among these constructs that account for their varying effects on mental and physical health (Watkins, 2008). Multidimensional scaling (MDS) of batteries of RT questionnaires identified two underlying dimensions that characterize RT constructs: valence and purpose. Valence contrasts RT with positive content (e.g., anticipating, reminiscing) against RT with negative content (e.g., rumination, self-criticism). Purpose contrasts searching, questioning, and uncertainty (e.g., reflection, self-analysis) against solving, certainty, and certainty-seeking (e.g., worry, planning). These valence and purpose dimensions, as well as an RT total score, emerge from batteries of structured RT questionnaires as well as evaluations of unstructured RT descriptions, in which respondents are asked to describe in a free-response format “what’s been ‘on your mind’, that is, you have thought about this topic frequently or for long periods of time” (Evans & Segerstrom, 2011; Segerstrom et al., 2003, 2010; Segerstrom, Stanton, Flynn, Roach, Testa, & Hardy, 2012).

The dimensional model provides a clearer representation of how and why RT correlates with other constructs, including psychological and physical health, than do models of discrete RT (Smith, McCarthy, & Zapolski, 2009). For example, older adults worry less than younger adults (Basevitz, Pushkar, Chaikelson, Conway, & Dalton, 2008). However, because worry is a combination of high total RT (i.e., the tendency to think repetitively), negative valence (i.e., the tendency to think about negative events), and solving purpose (i.e., the tendency to seek certainty and solve problems; Basevitz et al., 2008; Davey, 1994), it is difficult to know whether older adults think less repetitively, think more positively, or search more than younger adults. These possibilities have different implications, as RT valence has different and often larger correlations with health and well-being compared with purpose and total (Segerstrom et al., 2003, 2010, 2012; Watkins, 2008). In other examples, differentiating among RT dimensions revealed that more positive valence correlated positively with better executive function in older adults, whereas higher total RT correlated positively with higher estimated IQ (Segerstrom et al., 2010); more positive valence correlated positively with the non-judging aspect of mindfulness in young adults, whereas higher total RT correlated positively with the observing aspect (Evans & Segerstrom, 2011).

However theoretically desirable it would be to apply the dimensional model to questions about RT, a major barrier to applying the dimensional model is methodological. In previous studies, dimension scores have been derived from large batteries of RT questionnaires that can add to over 100 items (Evans & Segerstrom, 2011; Segerstrom et al., 2003, 2010).

Deriving the dimensions from unstructured RT descriptions is time-consuming both for the respondent and for coders and provides limited ability to assess total RT (Roach, Salt, & Segerstrom, 2010; Segerstrom et al., 2012). Where RT is the major focus of a study, the benefits of such a large questionnaire battery or of coding free-response data may outweigh the costs of participant and investigator burden. However, these measurement approaches are prohibitive when RT is not the major focus of a study or in methodologies with frequent repeated measurement (e.g., in a diary study). It is currently difficult for such studies to approach the measurement of RT from a dimensional perspective.

The goal of the current project was to develop an approach by which the "battery" of RT questionnaires required to derive dimension scores could be made more accessible by reducing it to a small number of items. Items in the ideal set would meet several criteria (Smith, McCarthy, & Anderson, 2000):

1. There should be few items in order to reduce burden as much as possible.
2. The items should positively correlate with each other and therefore capture the umbrella construct of RT and quantitative individual differences therein.
3. The items should adequately sample the multidimensional space and therefore capture qualitative individual differences in RT valence and purpose.
4. The items should demonstrate a balance between stability and change, neither idiosyncratic to a particular measurement point nor insensitive to change over time.

These criteria were applied to data from four samples that differed in age, geography, and method of data collection. Our objective was to identify items that would meet our criteria across samples and therefore demonstrate robustness to methodological differences. We then assessed the reliability, stability, and validity of these items with regard to the full battery method of assessment, other measures of RT, and depressive symptoms.

Method

Participants

Three archival samples and 3 samples collected specifically for this report contributed data. Table 1 includes their sources and demographic characteristics.

Procedure

Undergraduate (UG1, UG2, and UG3) samples completed questionnaires once in group or individual sessions for course credit. The UG4 sample completed up to 14 daily diaries online for course credit. Diaries were used for analysis if they were completed between 5 pm and 3 am on the designated day and the participant provided three or more valid diaries, resulting in 1,199 valid diaries from 99 participants. The online (WEB) sample completed questionnaires once on-line for personality feedback. The older adult (OA) sample completed questionnaires in individual interviews, with the assistance of an interviewer and

cards showing response options. In this longitudinal study, OA participants were assessed every six months for up to 13 waves. The full RT questionnaire battery was administered at Wave 1 and Wave 6. At Wave 11, the 75 remaining participants in the OA sample were joined by 73 new participants in a "refresher" sample. We refer to this refreshed sample as OA2. The brief item pool (see below) was administered to the OA2 participants at Waves 11, 12, and 13. OA participants received a gift card at each wave as thanks for their participation. All studies were performed with the approval of the University of Kentucky institutional review board.

Measures

Sources of candidate items—Candidate items from a diverse set of RT scales were evaluated in 4 samples: UG1, UG2, WEB, and OA.

The *Emotional Approach Coping – Processing (EAC-P)* scale consists of 4 items tapping coping through attention to and reflection on emotions (Stanton, Kirk, Cameron, & Danoff-Burg, 2000, Table 1). The trait form of the questionnaire was administered, using the stem, "When I am under a lot of stress…." EAC-P items were completed by all four samples.

The *Penn State Worry Questionnaire (PSWQ)* consists of 16 items tapping the propensity for pervasive, uncontrollable worry (Meyer, Miller, Metzger, & Borkovec, 1990, Table 1). PSWQ items were completed by all four samples.

The *Responses to Positive Affect Questionnaire (RPAQ)* consists of 17 items tapping repetitive thought about one's positive thoughts and emotions when happy (Feldman, Joormann, & Johnson, 2008, Table 1). RPAQ items were completed by the UG2 and WEB samples.

The rumination subscale of the *Response Style Questionnaire (RSQ*) consists of 22 items tapping repetitive thought about one's negative thoughts and emotions when sad, down, or depressed (Nolen-Hoeksema, Morrow, & Fredrickson, 1993; Treynor, Gonzalez, & Nolen-Hoeksema, 2003). RSQ items were completed by all four samples. Most items can be found in Treynor et al. (2003, Table 2). Two additional items from a different version (Nolen-Hoeksema et al., 1993) were also included: "When I feel sad, down, or depressed, I isolate myself and think about the reasons why" and "When I feel sad, down, or depressed, I try to understand myself by focusing on my depressed feelings" (S. Nolen-Hoeksema, personal communication, 4/15/96).

The *Rumination Scale (RS)* consists of 10 items tapping both rehearsal and uncontrollable repetitive thought (Martin, Tesser, & McIntosh, 1993). RS items were completed by the UG1, WEB, and OA samples. Three items were evaluated (see *Item Pool* below): "I often think about what my life will be like in the future" (item 4); "When I have a problem, I tend to think of it a lot of the time" (item 5); "When I do not understand something that happens, I tend to run it over in my mind until I can make sense of it" (item 10) (L.L. Martin, personal communication, 2/6/97).

The *Rumination-Reflection Questionnaire (RRQ)* consists of 12 items tapping rumination over threats to the self and 12 items tapping reflection on the self that is motivated by intellectual curiosity (Trapnell & Campbell, 1999, Table 3). RRQ items were completed by all four samples.

The *Savoring Beliefs Inventory (SBI* consists of 24 items tapping cognitive strategies for increasing pleasure from the past (i.e., reminiscing), present (i.e., savoring), and future (i.e., anticipating; Bryant, 2003). Six items were evaluated (see *Item Pool* below): "I feel a joy of anticipation when I think about upcoming good things" (item 7); "I can make myself feel good by remembering pleasant events from my past" (item 9); "When something good happens, I can make my enjoyment of it last longer by thinking or doing certain things" (item 11); "I like to store memories of fun times that I go through so that I can recall them later" (item 15); "I can make myself feel good by imagining what a happy time that is about to happen will be like" (item 19); "It's easy for me to rekindle the joy from pleasant memories" (item 21) (F. Bryant, personal communication, 2/27/01). SBI items were completed by the UG2 and WEB samples. A majority of the OA sample also completed the SBI items (n = 150). For the minority that did not (n = 29), to make maximal use of the OA data, these data were imputed using the expectation-maximization algorithm in SPSS and all of the other OA RT item responses as the basis for imputation (Schafer & Graham, 2002).

Response scales—UG1 and UG2 samples responded to RT items using the original response options for the individual scales. UG3, UG4, WEB, and OA samples responded to all items using a single 5-option response scale with options "strongly disagree", "disagree", "neutral", "agree", and "strongly agree" and the instruction to "indicate how accurately each of the statements describes you" for trait administration or "indicate how accurately each of the statements describes your thoughts today" for diary administration. We recommend that these instructions and response options be used in the future with the brief dimensional method we describe here. We converted all item responses to Z scores for analysis, which mitigates concern about scaling differences: Standardizing the responses would largely preserve the underlying distribution.

Validity measures—Validity of the dimension scores was assessed using two other measures of RT and a measure of depressive symptoms. Although RT dimensions have been associated with psychological well-being (Segerstrom et al., 2010; Segerstrom, Eisenlohr-Moul, Evans, & Ram, in press), negative affect generally and depressive symptoms specifically are a major focus of investigation in RT research (for example, see Aldao, Nolen-Hoeksema, & Schweizer, 2010, Watkins, 2008, for reviews). These validity analyses employed the UG3 and UG4 samples.

Pain catastrophizing is generally defined as "an exaggerated negative orientation toward noxious stimuli" (Sullivan, Bishop, & Pivik, 1995, p. 524). The *Pain Catastrophizing Scale* (PCS; Sullivan et al., 1995) contains three subscales: Rumination (e.g., "I keep thinking about how much it hurts"), Magnification (e.g., "I become afraid that the pain may get worse"), and Helplessness (e.g., "I feel I can't go on"). Alpha reliabilities for the subscales in UG3, who completed the scale, were .88, .61, and .84, respectively.

Unconstructive repetitive thought was measured daily in the UG4 sample with a 4-item scale designed to measure RT with negative content. Items refer to "personal problems and worries," thought that is "difficult to get out of your head," preoccupation with the future, and "situations that upset you" (M. Sliwinski, personal communication, 1/12/2014). In a validation study (M. Sliwinski, personal communication, 4/1/2015), the scale had good reliability both between (.99) and within (.74) people. Validation of the "unconstructive" label was tested with a multilevel confirmatory factor analysis. In this model, the unconstructive RT latent factor was distinct from positive and negative affect latent factors but correlated with both in the expected direction both between and within people (model RMSEA = .02). In the present study, using equations provided by Cranford and colleagues (2006), the scale reliability at Level 2 (between people across all days) was .99; at Level 1 (within people between days), .81.

Depressive symptoms were measured daily in the UG4 sample with the 10-item version of the *Center for Epidemiologic Studies – Depression (CESD)* scale (Kohout, Berkman, Evans, & Cornini-Huntley, 1993). The scale reliability at Level 2 (between people across all days) was .97; at Level 1 (within people between days), .73.

Data analysis

Multidimensional scaling (MDS)—MDS is a method for assessing the similarity of entities (e.g., items, scales, or people) along multiple dimensions, in contrast with other methods that cluster entities together. For example, a common goal of factor analysis is simple structure, in which entities load on one factor but not any others. MDS is suitable for alternatives to simple structure, including the circumplex structure that characterizes RT (Evans & Segerstrom, 2011; Segerstrom et al., 2003, 2010). Candidate items were standardized to Z scores and subjected to MDS using SPSS, with the ALSCAL method and ordinal distances and specifying 2 orthogonal dimensions.

Missing data—For the diary study (UG4), participants were given a sixth response option for items B, D, and E (see Figure 1). Although it is reasonable when inquiring about longer time frames to assume that people have felt happy, sad, or stressed at some point, it is possible that the same might not be true for a single day. Therefore, participants were given the option to say that they did not feel that state on that day. For item B (sad, down, or depressed) and item D (stress), the option was selected in 20% of diaries and for item E (happy) in 3% of diaries. All participants reported feeling each state on at least one diary day, with the exception of one person who did not report stress on any diary day. These responses were treated as missing data and imputed using the SPSS expectation-maximization algorithm with all other RT items as the basis for imputation (Schafer & Graham, 2002).

Multi-level models—For the diary study (UG4), SAS PROC MIXED with maximum likelihood estimation was used. Days were at Level 1 and people at Level 2. Intraclass correlations (ICCs) were calculated from variance estimates produced by a null model with no predictors. In the models predicting unconstructive RT and depressive symptoms from RT dimensions, RT dimension predictors were separated into between-person, Level 2

predictors (person means, grand-mean centered) and within-person, Level 1 predictors (day deviations from person means; Enders & Tofighi, 2007). RT predictors were standardized so that intervals of 1 equaled 1 SD difference in the predictor. Inclusion of Level 1 random effects was determined by comparison of the −2 log likelihood in models with and without the random effect.

Results

Item selection

Candidate items—Item selection was based on candidate item data from UG1, UG2, WEB, and OA samples (N = 1,588). From the RT questionnaires described above, candidate items were selected that were likely to (1) sample the multidimensional space adequately, based on the multidimensional position of the total questionnaire, (2) demonstrate adequate item-total correlations, based on questionnaire-level correlations, and (3) have adequate test-retest reliability while also having the possibility for change. This third criterion was applied because descriptions of RT show both stability and change in valence and purpose (Roach et al., 2010). Items were examined and eliminated for explicit trait content (e.g., “I have been a worrier all my life” from the PSWQ) and preferred when correlations in the OA sample between Waves 1 and 6 were in the range of .30 – .50. This process resulted in a pool of 46 candidate items.

MDS—A two-dimensional model produced the valence and purpose dimensions in each sample and accounted for 92% of the variance in UG1, 83% in UG2, 87% in WEB, and 80% in OA. These percentages are consistent with results of MDS on batteries of RT scale scores in UG1, UG2, and OA samples (Evans & Segerstrom, 2011; Segerstrom et al., 2003, 2010).

Across the four models, individual-model and mean item positions and item-total correlations[1] were used to select items that represented eight equal and symmetric sectors of the multidimensional space (see Figure 1). These items successfully created a circumplex structure. Note that this is the minimum number of items needed to create a symmetrical circumplex (e.g., 4 items would not create a circular structure, and 6 items would not be perfectly symmetrical). Initial selection of some reverse-scored items resulted in poor item-total correlations in the 8-item subset, and so these were replaced with items that did not require reverse scoring. In Figure 1, sectors are identified by letters, and the diamonds in Figure 1 show the average position of each selected item across all samples. The position of each item is generally consistent with the position of the questionnaire from which it was drawn in studies using scale batteries (Evans & Segerstrom, 2011; Segerstrom et al., 2003, 2010). Correlations among the items were consistent with a circumplex structure (Gurtman & Pincus, 2003): The mean of the correlations between adjacent items (converting to Z to take the mean, and using all available data) was .25; with one separation, .16; with two separations, .07; and for items directly across from each other, −.01.

[1]A table with the MDS results for all items in all samples is available by request from the first author.

Brief assessment of RT dimensions

Scoring—Exact dimensional positions from MDS have been used to weight questionnaire scores to yield dimension scores in previous studies (e.g., Evans & Segerstrom, 2011; Segerstrom et al., 2010). In order to avoid over-fitting to our samples and maximize generalizability of the results to other samples, we weighted items according to their target symmetrical circumplex position for the valence and purpose dimensions (i.e., the cosine of the angle between the target position and the appropriate dimension axis; Strack, Jacobs, & Holtforth, 2013) and weighted them evenly for the total dimension. Users of this approach could alternatively determine the exact multidimensional positions of the items in their sample(s) and use those positions to weight standardized item scores in calculating dimension scores. The calculation of the dimension scores was as follows, noting that these calculations should be applied to the standardized items, even when a single response scale is used:

Valence = 0.92388*(A+H) + 0.38268*(G+B) – 0.38268*(C+F) – 0.92388*(E +D)

Purpose = 0.92388*(B+C) + 0.38268*(D+A) – 0.38268*(E+H) – 0.92388*(F +G)

Total = sum (A through H)

Higher scores on the valence dimension reflect more negative content, and higher scores on the purpose dimension reflect more searching.

Reliability and stability—Axis reliability was assessed using a CFA method for assessing the reliability of circumplex axes (Strack et al., 2013) implemented in Mplus (Muthén & Muthén, 1998–2012). UG3, WEB, and OA2 (Wave 11) samples (N = 493) completed all 8 items using the same instructions and response scale. Therefore, these data were combined for this analysis to provide the maximum sample size as well as representation across the adult lifespan. The general factor (total) and axes (valence and purpose) were modeled with item weights of 1.0 on the general factor and weights corresponding to the ideal circumplex position on the axes (see Scoring above). The model fit well (CFI = .90, rmsea = .049 [CI = .033 – .066]). Reliability of the total score was .61; reliability of the valence and purpose axes was .52. These reliabilities are consistent with many other circumplex measures and higher than other 8-item circumplex measures (Strack et al., 2013).

The OA2 and UG4 samples completed the items multiple times; therefore, dimension stability (i.e., test-retest reliability) can also be estimated. ICCs were employed to indicate the dimension variance shared across administrations. In the OA2 sample, with “trait” administration, ICCs were .49 for valence, .42 for purpose, and .70 for total across up to 3 waves at 6-month intervals (i.e., over a 1-year period). In the UG4 sample with “today” administration, ICCs were .32 for valence, .26 for purpose, and .48 for total across up to 14 days at 1-day intervals (i.e., over a 2-week period). Therefore, the brief assessment items met our criterion for having adequate stability while still allowing for the possibility of change.

Validity—A subset of the OA2 sample (N = 75) came from the OA sample, permitting the calculation of correlations between their brief assessment dimension scores (averaged across Waves 11, 12, and 13) and their dimension scores derived from a full RT battery completed 5–6 years previously (Segerstrom et al., 2010). The correlation for valence was .50 ($p < .001$); for purpose, .34 ($p < .003$); and for total, .48 ($p < .001$). None of the off-diagonal correlations exceeded .20, and none was statistically significant. Therefore, the dimension scores derived from 8 items not only correlated with the full-battery dimension scores, suggesting that they are a valid representation of full-battery scores, but did so over a long period of time.

The next validity evidence derives from correlations between the brief assessment dimension scores and other rumination measures. Correlations with the Pain Catastrophizing Scale Rumination subscale provided evidence of convergent validity, whereas correlations with the Magnification and Helplessness subscales provided evidence of divergent validity. Table 2 shows the correlations between RT dimensions and these subscales in the UG3 sample. The propensity to ruminate over pain was associated with a propensity for engaging in more and more negatively valenced RT but not strongly associated with a propensity for searching or solving RT. Correlations between RT dimensions and Magnification and Helplessness were computed partialing the Rumination score, as Rumination was highly correlated with the other two subscales ($r = .62–.64$). The correlations between RT dimensions and Magnification and Helplessness were small and not statistically significant after accounting for their overlap with Rumination.

Table 3 shows the relationships between RT dimensions and unconstructive RT in the diary study (UG4). At Level 2 (between people), participants who generally tended to engage in more and more negative RT also tended to report more unconstructive RT across days; however, RT purpose was unrelated to unconstructive RT at this level. At Level 1 (within people), on days when RT was higher, more negative, and more searching, participants also reported more unconstructive RT. There were random effects of RT dimensions on unconstructive RT; that is to say, the within-person relationship varied across participants. For example, for some people, unconstructive RT was higher on days when they had more negatively valenced RT; for others, unconstructive RT was higher on days when they had more total RT. The range of estimated slopes across the sample is shown in Table 3.

Final validity evidence is provided by the relationships between the RT dimensions and daily reports of depressive symptoms in the diary study (UG4; see Table 3). At Level 2 (between-person), participants who typically engaged in more negative RT had more depressive symptoms across the diary period. To a lesser degree, participants who typically engaged in more total RT had more depressive symptoms across the diary period. RT purpose was not significantly related to depressive symptoms at this level. At Level 1 (within-person), more RT, more negative RT, and more searching RT on a particular day all associated with more depressive symptoms on that day. RT valence had the strongest relationship to depressive symptoms; however, the magnitudes of the dimension effects were more equal at Level 1 (daily changes) than at Level 2 (individual differences). There were also random effects of valence and total, indicating that for some individuals, these RT dimensions had stronger relationships with depressive symptoms than for other individuals.

Discussion

Multidimensional scaling of large batteries of RT questionnaires yields three underlying dimensions: valence, purpose, and total RT (Evans & Segerstrom, 2011; Segerstrom et al., 2003, 2010). However, the full-battery measurement method may impose a prohibitively large burden on participants in many studies, including studies with many other questionnaires, with repeated measures, or with special populations. The present analyses were intended to identify a small number of RT items that could stand in for the large battery of questionnaires currently required to capture the qualitative and quantitative dimensions that characterize essential properties of RT.

These items were selected to create a circumplex structure based on their empirical positions in multidimensional space in 4 samples representing the adult lifespan (ages 18–94 years). In general, the items were face valid, with items suggesting negative thought content at the negative valence pole (e.g., item A) and items suggesting positive thought content at the positive pole (e.g., item F). Likewise, items suggesting uncertain or questioning content were located at the searching pole (e.g., item C), and items suggesting certain or planning content were located at the solving pole (e.g., item G). One possible exception to this face validity is item D, which makes explicit reference to acknowledging one's emotions under stress. However, other evidence supports the empirical finding that this kind of RT belongs on the positive pole of the valence dimension. First, the emotional processing scale has consistently been located at the positive pole in questionnaire-level analyses (Evans & Segerstrom, 2011; Segerstrom et al., 2003, 2010). Second, women with higher scores on the emotional processing scale had more explicitly positive thought content in unstructured RT descriptions (Segerstrom et al., 2012), indicating that engaging in emotional processing either follows from or facilitates more positive RT content. Finally, both of the most positive RT items (D and E) involve emotional approach, which is associated with better emotion regulation, more positive and less negative affect under stress, and more effective use of emotion as an adaptive guide for behavior (Austenfeld & Stanton, 2004; Beals, Peplau, & Gable, 2009; Kashdan, Barrios, Forsyth, & Steger, 2006; Mennin, Heimberg, Turk, & Fresco, 2002; Stanton et al., 2000).

One could argue that the valence dimension could also be labeled "adaptiveness" or "control". The valence label was chosen to reflect content differences and was validated against positive and negative content in unstructured descriptions (Segerstrom et al., 2010). Although the valence dimension of RT does correlate with adaptive outcomes and perceived control over thoughts, such a label presumes that positively valenced thoughts are always adaptive or controllable, which may not be the case. We view valence as the label that presumes the least about empirically testable correlates of the dimension.

There were four criteria for the assessment items. First, there should be few items. The minimum number of items required to form a symmetric circumplex were selected, maximizing the utility of the scale for applications in which long assessments of RT are impractical. Second, the items should positively correlate with each other, capturing individual differences in total RT, that is, the tendency to think repetitively. The success of these items in doing so is reflected primarily by the adequate reliability of the total RT

dimension, but also by the finding that items opposite each other in the circumplex were not negatively correlated but rather uncorrelated. Therefore, item association ranged from neutral to positive. Third, the items should adequately sample the multidimensional space and capture valence and purpose. The selected items formed a circumplex and provided adequate axis reliability (and higher reliability than comparable measures; Strack et al., 2013). There was the expected trade-off between scale length and reliability, such that axis reliability was adequate but not ideal. Smith et al. (2001) suggest that short forms should demonstrate adequate time savings for any loss of reliability and validity: In this case, assessment time is cut by 92% when administering these 8 items compared with the 108 items for the full battery of RT measures described above. This method of assessing RT dimensions therefore makes such assessment possible for screening purposes, intensive data collection (e.g., in diaries or ecological momentary assessment), or large-scale surveys (Robins, Hendin, & Trzesniewski, 2001). Finally, the items should demonstrate a balance between stability and change. For the valence dimension, 49% of the variance was attributable to stable individual differences over 1 year with instructions to “indicate how accurately each of the statements describes you”, and 32% over 2 weeks with instructions to “indicate how accurately each of the statements describes your thoughts today”. The corresponding proportions for the purpose dimension were 42% and 26%, and for the total dimension, 70% and 48%. Total amount of RT therefore showed the highest test-retest stability, followed by valence and then purpose. This pattern is consistent with the degree of stability and change in unstructured descriptions of RT, in which rater-assessed RT valence was more stable within people than was rater-assessed purpose (Roach et al., 2010). Therefore, these items seem to reflect the reality of RT in their degree of stability and change over time. Stability is further demonstrated by the substantial correlations between dimension scores derived from the full battery of RT measures and the dimension scores derived from the 8 RT items 5–6 years later. In the diary study, stable individual differences and changes in RT dimensions were both associated with depressive symptoms. These associations indicate the importance of both levels of variance in RT and the success of the items in capturing both stability and change.

The validity evidence demonstrates some uses of the brief assessment of RT dimensions. First, correlating individual RT constructs with the underlying dimension scores allows for location of the scale in the RT dimensional space and the RT nomological net. A failure to obtain a positive correlation with RT total might lead one to question whether the scale belongs in the RT nomological net at all. If a positive correlation is obtained, correlations with RT valence and purpose can be used to estimate its similarity and difference to other constructs without having to administer any further RT questionnaires. For example, rumination in the context of pain catastrophizing correlated .38 with RT total, confirming its membership in RT. Figure 2 shows the location of this scale in multidimensional space (determined by first correcting correlations in Table 2 for attenuation and then converting correlations to angles). Other RT located in this space includes worry (from the Penn State Worry Questionnaire) and rumination (from the Rumination-Reflection Questionnaire; Segerstrom et al., 2003). Unconstructive RT was also significantly correlated with RT total and can likewise be located in multidimensional space. Between people, unconstructive RT is nomologically similar to lack of control over thoughts (from the Rumination Scale) and

self-reproach (from the Response Styles Questionnaire; Segerstrom et al., 2003). Note that unconstructive RT location was different at the between-person level and the within-person level insofar as increases in unconstructive RT were associated with even higher searching purpose at the within-person level.

One rationale for the dimensional model of RT is that most RT constructs are complex combinations of valence, purpose, and total, making it difficult to attribute RT correlates to a particular RT quality. In contrast, brief assessment dimension scores permit such attributions. The diary study also demonstrated that the qualities underlying a measure of unconstructive RT varied across people. For some people, an increase in unconstructive RT reflected a day on which they engaged in more negative RT; for others, a day on which they engaged in more RT of all kinds (not just the negative kind). Insofar as the dimensions have different relationships to outcomes, it is informative to know what underlying process is responsible for variations in unconstructive RT. To return to the earlier example, age differences in worry may result from age differences in valence, purpose, total, or some combination thereof. Understanding these underlying age differences as well as potential individual differences within age groups (e.g., some older adults may engage in less total RT than their younger counterparts, whereas others may engage in more positively valenced RT) would contribute to a more nuanced understanding of changes in RT with age.

Brief assessment of RT dimensions permits use in diary studies, which the full-battery method does not. Diary study results demonstrate the utility of the dimensional method. As expected, more negatively valenced RT was most highly associated with more depressive symptoms both between people and within people. However, higher total and more searching purpose were also associated with depressive symptoms, relationships that are not always observed in “trait”, cross-sectional designs (e.g., Segerstrom et al., 2010); more searching purpose was associated with depressive symptoms only within people. The difference may result from differences in diary vs. retrospective report (e.g., Stone et al., 1998; Todd, Tennen, Carney, Armeli, & Affleck, 2004). Notably, this multidimensional position (high negative searching) is characteristic of depressive rumination, with “brooding” typically in the A octant and “reflection” typically in the B octant (see Figure 1; Segerstrom et al., 2003; Treynor et al., 2003). The diary methodology also reveals interesting individual differences in the relationship of RT to depressive symptoms. Some individuals have strong relationships between negatively valenced thinking and depressive symptoms, whereas for others, the valence of their thoughts appeared to have minimal effect. These individual differences are an important area for future investigation.

The major limitation of the brief assessment of RT dimensions was related to its use in a daily diary: In a minority of diaries, people reported not experiencing sadness, stress, or happiness on a given day. These states were nearly universal on a weekly basis (i.e., all but one person reported each of the states at some point during the week), so this method of assessing RT can reasonably be used without an opt-out option for items B, D, and E (e.g., “I did not feel happy at any time today”) for time scales of a week or longer. Where people do opt out of these items, because each item has a structural relationship to the other items, it is reasonable to impute those values. Nonetheless, the necessity of doing so is a limitation of this method. A second limitation of the data concerns diversity of the samples. The evidence

presented in this report was derived from samples that varied in age, data collection method, and geography, attesting to the generalizability of the findings. However, only two samples were ethnically diverse (UG1, WEB). Therefore, more work should provide reliability and validity evidence in diverse samples.

Conclusion

A dimensional approach to the assessment of RT makes it possible to pose and answer important questions about how RT constructs relate to each other (e.g., where does unconstructive RT fall in the nomological net?) and why they relate to other constructs (e.g., which dimension most strongly correlates with depressive symptoms?) The greater accessibility of dimensional assessment of RT using this smaller number of items should facilitate future work to address important questions about the qualities of RT that predict mental and physical health.

Acknowledgments

This research was supported by NIH [AG026307-R01, AG039208-F31, AG048692-F31, and AG033629-K02]. The authors thank Gregory T. Smith for his comments on an earlier version of this work and Martin Sliwinski for provision of the unconstructive repetitive thought items and validity evidence.

References

Aldao A, Nolen-Hoeksema S, Schweizer S. Emotion-regulation strategies across psychopathology: A meta-analytic review. Clinical Psychology Review. 2010; 30:217–237. [PubMed: 20015584]

Austenfeld JL, Stanton AL. Coping through emotional approach: a new look at emotion, coping, and health-related outcomes. Journal of Personality. 2004; 72:1335–1364. [PubMed: 15509285]

Basevitz P, Pushkar D, Chaikelson J, Conway M, Dalton C. Age-related differences in worry and related processes. International Journal of Aging and Human Development. 2008; 66:283–305. [PubMed: 18507331]

Beals KP, Peplau LA, Gable SL. Stigma management and well-being: The role of perceived social support, emotional processing, and suppression. Personality and Social Psychology Bulletin. 2009; 35:867–879. [PubMed: 19403792]

Bryant FB. Savoring Beliefs Inventory (SBI): A scale for measuring beliefs about savouring. Journal of Mental Health. 2003; 12:175–196.

Cranford JA, Shrout PE, Iida M, Rafaeli E, Yip T, Bolger N. A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? Personality and Social Psychology Bulletin. 2006; 32:917–929. [PubMed: 16738025]

Davey, GCL. Pathological worrying as exacerbated problem-solving. In: Davey, GCL.; Tallis, F., editors. Worrying: Perspectives on Theory, Assessment, and Treatment. New York: Wiley; 1994. p. 35-59.

Enders CK, Tofighi D. Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. Psychological Methods. 2007; 12:121–138. [PubMed: 17563168]

Evans DR, Segerstrom SC. Why do mindful people worry less? Cognitive Therapy and Research. 2011; 35:505–510.

Feldman GC, Joormann J, Johnson SL. Responses to positive affect: A self-report measure of rumination and dampening. Cognitive Therapy and Research. 2008; 32:507–525. [PubMed: 20360998]

Gurtman, MB.; Pincus, AL. The circumplex model: Methods and research applications. In: Schinka, JA.; Velicer, WF., editors. Handbook of Psychology: Volume 2, Research Methods in Psychology. Hoboken, NJ: Wiley; 2003. p. 407-427.

Kashdan TB, Barrios V, Forsyth JP, Steger MF. Experiential avoidance as a generalized psychological vulnerability: Comparisons with coping and emotion regulation strategies. Behaviour Research and Therapy. 2006; 44:1301–1320. [PubMed: 16321362]

Kohout FJ, Berkman LF, Evans DA, Cornoni-Huntley J. Two shorter forms of the CES-D depression symptoms index. Journal of Aging and Health. 1993; 5:179–193. [PubMed: 10125443]

Martin, LL.; Tesser, A.; McIntosh, D. Wanting but not having: The effects of unattained goals on thoughts and feelings. In: Wegner, DM.; Pennebaker, JW., editors. Handbook of Mental Control. Englewood Cliffs, NJ: Prentice Hall; 1993.

Mennin DS, Heimberg RG, Turk CL, Fresco DM. Applying an emotion regulation framework to integrative approaches to generalized anxiety disorder. Clinical Psychology: Science and Practice. 2002; 9:85–90.

Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the Penn State Worry Questionnaire. Behaviour Research and Therapy. 1990; 28:487–495. [PubMed: 2076086]

Nolen-Hoeksema S, Morrow J, Fredrickson BL. Response styles and the duration of episodes of depressed mood. Journal of Abnormal Psychology. 1993; 102:20–28. [PubMed: 8436695]

Roach AR, Salt CE, Segerstrom SC. Generalizability of repetitive thought: Examining stability of thought content and process. Cognitive Therapy and Research. 2010; 34:144–158.

Robins RW, Hendin HM, Trzesniewski KH. Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. Personality and Social Psychology Bulletin. 2001; 27:151–161.

Schafer JL, Graham JW. Missing data: Our view of the state of the art. Psychological Methods. 2002; 7:147–177. [PubMed: 12090408]

Segerstrom SC, Eisenlohr-Moul TA, Evans DR, Ram N. Repetitive thought dimensions, psychological well-being and perceived growth in older adults: A multilevel, prospective study. Anxiety, Stress, and Coping. (in press).

Segerstrom SC, Roach AR, Evans DR, Schipper LJ, Darville AK. The structure and health correlates of trait repetitive thought in older adults. Psychology and Aging. 2010; 25:505–515. [PubMed: 20677888]

Segerstrom SC, Stanton AL, Alden LE, Shortridge BE. A multidimensional structure for repetitive thought: What's on your mind, and how, and how much? Journal of Personality and Social Psychology. 2003; 85:909–921. [PubMed: 14599253]

Segerstrom SC, Stanton AL, Flynn SM, Roach AR, Testa JJ, Hardy JK. Episodic repetitive thought: Dimensions, correlates, and consequences. Anxiety, Stress, and Coping. 2012; 25:3–21.

Smith GT, McCarthy DM, Anderson KG. On the sins of short-form development. Psychological Assessment. 2000; 12:102–111. [PubMed: 10752369]

Smith GT, McCarthy DM, Zapolski TC. On the value of homogeneous constructs for construct validation, theory testing, and the description of psychopathology. Psychological Assessment. 2009; 21:272–284. [PubMed: 19719340]

Stanton AL, Kirk SB, Cameron CL, Danoff-Burg S. Coping through emotional approach: Scale construction and validation. Journal of Personality and Social Psychology. 2000; 78:1150–1169. [PubMed: 10870915]

Stone AA, Schwartz JE, Neale JM, Shiffman S, Marco CA, Hickcox M, Cruise LJ. A comparison of coping assessed by ecological momentary assessment and retrospective recall. Journal of Personality and Social Psychology. 1998; 74:1670–1680. [PubMed: 9654765]

Strack M, Jacobs I, Holtforth MG. Reliability of circumplex axes. SAGE Open. 2013; 3 2158244013486115.

Sullivan MJ, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychological Assessment. 1995; 7:524–532.

Todd M, Tennen H, Carney MA, Armeli S, Affleck G. Do we know how we cope? Relating daily coping reports to global and time-limited retrospective assessments. Journal of Personality and Social Psychology. 2004; 86:310–319. [PubMed: 14769086]

Trapnell PD, Campbell JD. Private self-consciousness and the five-factor model of personality: Distinguishing rumination from reflection. Journal of Personality and Social Psychology. 1999; 76:284–304. [PubMed: 10074710]

Treynor W, Gonzalez R, Nolen-Hoeksema S. Rumination reconsidered: A psychometric analysis. Cognitive Therapy and Research. 2003; 27:247–259.

Watkins ER. Constructive and unconstructive repetitive thought. Psychological Bulletin. 2008; 134:163–206. [PubMed: 18298268]

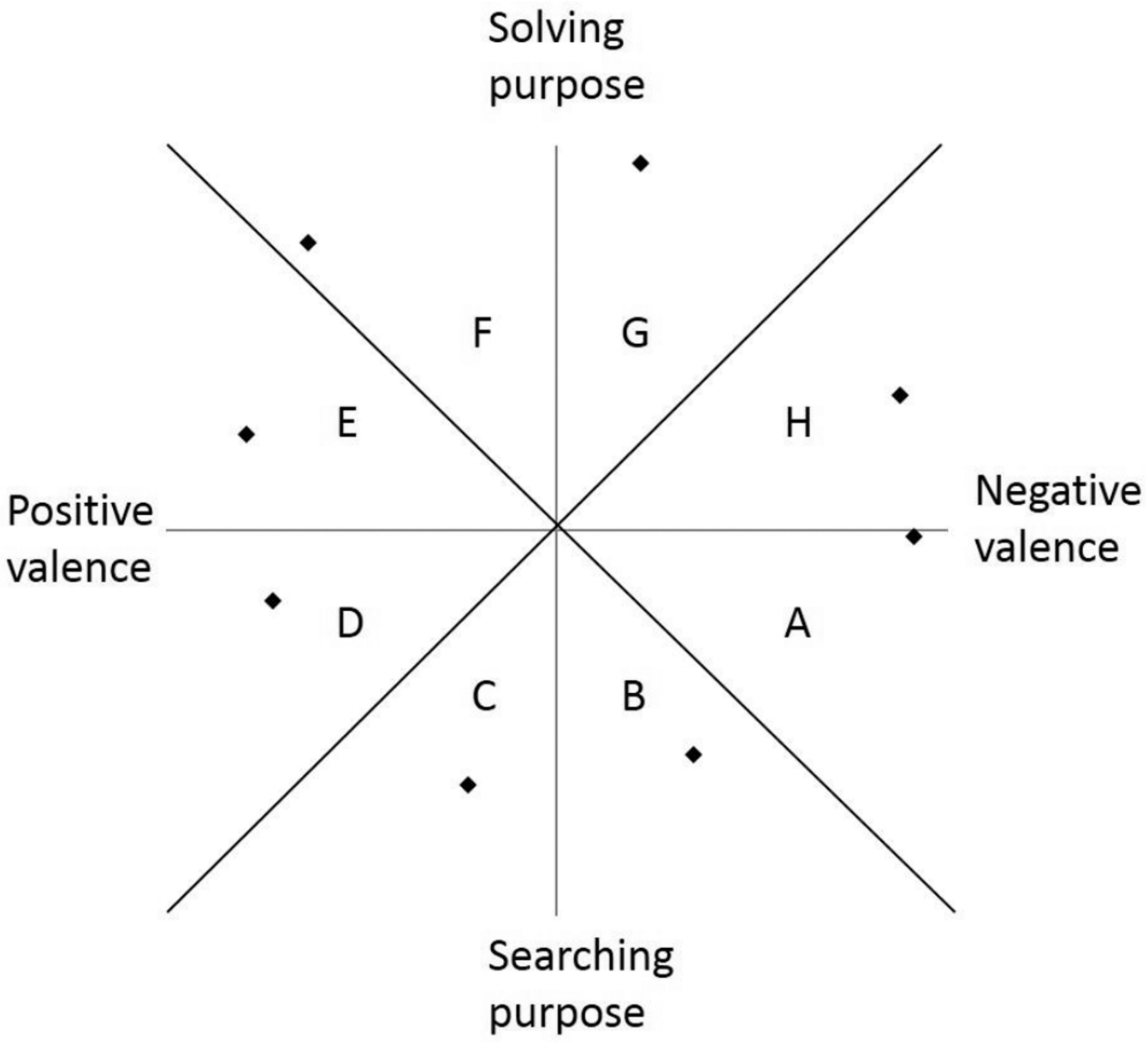

Figure 1.

The multidimensional RT space and mean location of selected items for each of 8 sectors across the 4 samples (UG1, UG2, WEB, OA)

A: I spend a great deal of time thinking back over my embarrassing or disappointing moments.

B: When I feel sad, down, or depressed, I isolate myself and think about the reasons why.

C: I love to meditate on the nature and meaning of things.

D: When I am under a lot of stress, I acknowledge my emotions.

E: When I feel happy, I note how I feel full of energy.

F: I like to store memories of fun times that I go through so that I can recall them later.

G: I often think about what my life will be like in the future.

H: I know I should not worry about things, but I just cannot help it.

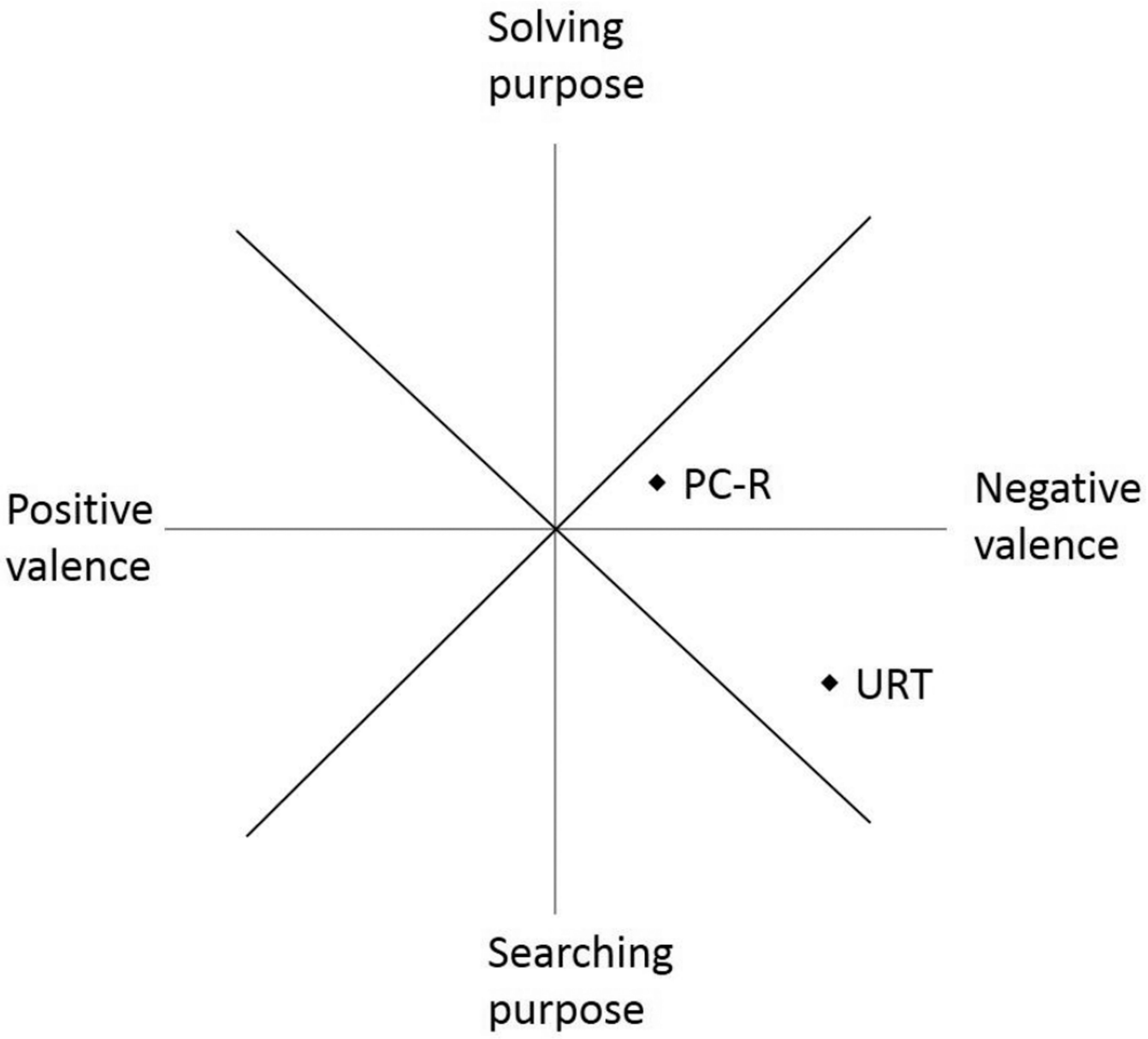

Figure 2.
Location of rumination in pain catastrophizing (PC-R) and unconstructive repetitive thought (URT) in dimensional space, based on between-subjects correlations with RT dimensions

Table 1

Participant characteristics

Sample	N	M Age (SD)	% Caucasian	% Female	Source
UG1	954	18.9 (1.8)	57	64	Segerstrom et al., 2003
UG2	199	*	*	*	Evans & Segerstrom, 2011
UG3	113	19.5 (2.4)	*	49	-
UG4	99	*	*	72	-
WEB	232	36.8 (15)	68	78	-
OA	179	75.0 (6.5)	96	60	Segerstrom et al., 2010
OA2[a]	148	75.4 (5.9)	94	58	-

Note: UG = undergraduate; OA = older adults

[*] Demographic data were not collected, but the typical participant pool from which the sample was drawn was mean age of 19 years, 63% female, and identified ethnicities as White or Caucasian (87%), Black or African-American (8%), or other (5%).

[a] Includes N = 75 from the OA sample; see *Procedure*.

Table 2

Correlations between RT dimensions and aspects of pain catastrophizing (UG3 sample)

	Valence	Purpose	PCS Rumination	PCS Magnifying[a]	PCS Helplessness[a]
Valence	-		.27*	.12	.14
Purpose	−.02	-	−.12	.04	−.11
Total	.05	−.08	.38*	−.04	.06

Note. PCS = Pain Catastrophizing Scale. Higher scores on valence are more negative; higher scores on purpose are more searching.

*$p < .05$

[a]Partial correlations accounting for overlap with Rumination.

Assessment. Author manuscript; available in PMC 2017 October 01.

Table 3

Relationships between RT dimensions and unconstructive RT and depressive symptoms in the diary study (UG4 sample)

	Unconstructive RT (range = 0 – 100)			**Depressive symptoms (range = 0 – 10)**		
Fixed effects	Effect	SE	Slope Range	Effect	SE	Slope Range
Intercept	39.88*	1.25	-	2.29*	0.13	-
Valence (between)	9.41*	1.32	-	0.83*	0.14	-
Purpose (between)	1.57	1.38	-	0.24	0.14	-
Total (between)	8.96*	1.28	-	0.31*	0.13	-
Valence (within)	4.63*	0.71	−2.59 to 11.30	0.47*	0.07	−0.02 to 1.00
Purpose (within)	1.99*	0.64	−2.65 to 6.40	0.39*	0.06	-
Total (within)	5.61*	0.71	0.63 to 13.70	0.25*	0.07	−0.29 to 0.97
Random effects	Estimate	SE		Estimate	SE	
Intercept variance	130.03*	22.55		1.29*	0.23	
Valence slope variance	16.23*	6.58		0.12*	0.06	
Purpose slope variance	10.41*	5.28		-	-	
Total slope variance	15.56*	7.11		0.14*	0.07	
Residual variance	178.34*	11.11		2.32*	0.13	

Note. Higher scores on valence are more negative; higher scores on purpose are more searching.

*$p < .05$